

**MODELS MIXTOS MULTINIVELL EN
L'ESTUDI DE L'EVOLUCIÓ DEL DEFECTE
ÒPTIC A NENS AMB SINDROME DE DOWN**

M. Teresa Bonet
Imma Aibar

Director del projecte: Josep Anton Sànchez
Juny 2008

1. Introducció	1
1.1. Síndrome de Down	1
1.2. L'ull	6
2. Objectius	18
3. Materials i mètodes	21
3.1. Formulació del model.....	22
3.2. Estimació dels paràmetres.....	25
3.3. Inferència	27
3.4. Aspectes tècnics (computacionals)	31
3.5. Referència a la llibreria utilitzada	39
4. Resultats	54
4.1. Base de dades.....	54
4.2. Descriptius	57
4.3. Model Lineal Simple	66
4.4. Models lineals mixtos.....	71
4.5. Models lineals mixtos sense atípics.....	93
4.6. Models lineals mixtos sense atípics afegint covariables .	98
4.7. Resum dels models.....	104
4.8. Interpretacions.....	106
5. Conclusions	108
5.1. Possibles extensions	109
6. Bibliografia	110
7. Annex	113
7.1. Sintaxi utilitzada al programa SPSS (versió 15.0)	113
7.2. Sintaxi utilitzada al programa R (versió 2.3.0)	121
7.3. Modelització dels individus de la mostra	131
7.4. Modelització dels individus de la mostra sense atípics .	141

1. Introducció

La Fundació Catalana Síndrome de Down dona suport des del diagnòstic prenatal i naixement fins l'edat adulta, i els seus serveis estan adreçats a persones amb la síndrome de Down i altres discapacitats intel·lectuals. La Fundació rep la col·laboració de diferents professionals de la medicina. En concret la Dra. Alícia Galán del Servei d'oftalmologia del Hospital Vall d'Hebron fa els controls i seguiments oftalmològics dels nens amb Síndrome de Down. Com a resultat d'aquesta activitat, la doctora ha generat una base de dades amb les dades de les visites, que conté la informació que permet fer un seguiment de l'estat i evolució de la capacitat visual dels pacients. L'anàlisi estadística d'aquestes dades ha estat objecte de treball en comú del director d'aquest treball amb l'experta per tal de caracteritzar l'evolució dels errors refractius i la seva relació amb patologies i intervencions. La modelització de les dades en el context de les dades longitudinals mitjançant models mixtos ha estat l'objectiu marcat en el present treball.

1.1. Síndrome de Down

El material genètic de les persones està format per 46 cromosomes, dels quals 23 venen del pare i els altres 23 de la mare. Els parells de cromosomes són molt similars entre si, excepte en l'anomenat cromosoma sexual on el X és molt diferent de la Y, això únicament passa en el cas dels homes, tenen el parell sexual XY. En canvi, en les dones el parell de cromosomes sexuals són iguals, tenen el parell XX. Cada parell de cromosoma rep un nom, tenim el parell sexual i la resta estan enumerats de l'1 al 22. El Síndrome de Down (a partir d'ara SD), es caracteritza per un excés de material genètic corresponent al parell cromosomàtic 21. En aquest excés de material genètic es poden diferenciar 3 casos diferents:

- **Trisomia 21:** En aquest cas, l'individu presenta 3 cromosomes 21. (En total té 47 cromosomes). Aquest és el cas més habitual. Passa en el 95% dels casos.
- **Translocació:** En aquest cas, existeixen només 2 cromosomes 21, però tenim un fragment de material genètic del cromosoma 21 enganxat a algun altre

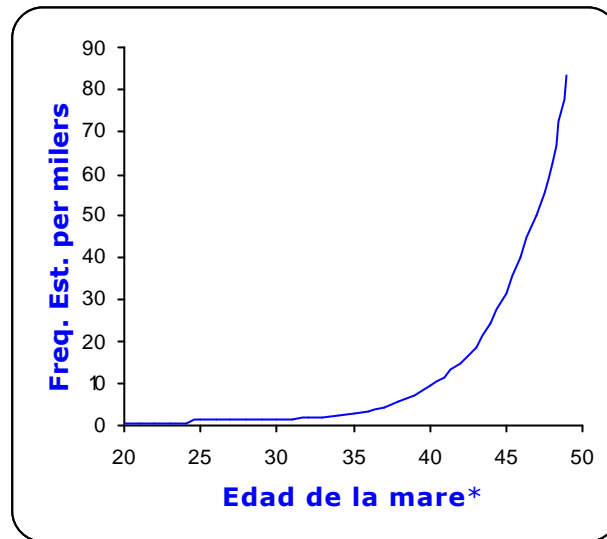
cromosoma. En aquest cas, tindríem un fragment del cromosoma 21 que existiria per triplicat. Aquest cas passa en un 4% dels casos.

- **Mosaïcisme:** En aquest cas, el problema es produeix durant la gestació. En algun moment de l'embaràs, alguna cèl·lula es duplica amb l'error de Trisomia 21. A partir d'aquest moment, cada vegada que la cèl·lula es multiplica, transmet l'error a les seves filles. El resultat és un organisme on conviuen cèl·lules trisòmiques amb cèl·lules normals. Això passa en un 1% dels casos de SD.

La incidència en la població en general és d'1 cas de cada 700. De totes maneres, la incidència general considerant el total de concepcions (infants nats vius i avortats) és més del doble d'aquesta xifra. Sabem que existeix una correlació entre l'edat de la mare en el moment del part i el problema del cromosoma 21. Aquest augment del risc en relació de l'edat de la mare en el moment del part està distribuïda de la següent manera:

Edat de la mare*	Freqüència estimada	Edat de la mare*	Freqüència estimada
20	1/1925	35	1/365
21	1/1695	36	1/285
22	1/1540	37	1/225
23	1/1410	38	1/175
24	1/1300	39	1/140
25	1/1205	40	1/110
26	1/1125	41	1/85
27	1/1050	42	1/67
28	1/990	43	1/53
29	1/935	44	1/41
30	1/885	45	1/32
31	1/825	46	1/25
32	1/725	47	1/20
33	1/590	48	1/16
34	1/465	9	1/12

* Edat de la mare en el moment del part



* Edat de la mare en el moment del part

Diagnòstic Prenatal

El diagnòstic durant l'embaràs consisteix en analitzar cèl·lules procedents de l'embrió, i fer un recompte dels cromosomes que porten. Aquesta visualització s'anomena cariotip. Existeixen dues alternatives:

- Biòpsia de Còrion: Es pot realitzar des de la setmana 10 a la 13 de l'embaràs.
- Amniocentesi: S'obté líquid amniòtic a partir d'una petita punxada a la membrana que envolta el fetus. Es pot dur a terme a partir de la setmana 15.

A més a més d'aquest diagnòstic, es poden dur a terme determinats anàlisis bioquímics de sang, que busquen la presència quantitativa de certes molècules que poden informar sobre la "possibilitat" o el "risc". Aquests anàlisis es poden dur a terme des del inici de l'embaràs.

Característiques

Podem dividir les característiques de les persones amb SD en 3 grups: nivell cognitiu, trets físics i fisiologia.

Nivell cognitiu:

Si agafem com a base els test de Coeficient Intel·lectual estàndard, freqüentment els nens amb SD són classificats amb nivells de retard mental dèbils, inclòs, cert percentatge es troba en grups amb coeficient normal-baix. Aquests test no són una referència única ja que sovint els nens sorprenen per la seva memòria, la intuïció, la creativitat,...

Els nens amb SD no estan condemnats a quedar-se amb un nivell intel·lectual equivalent a nens sense discapacitat de 5 a 8 anys. Actualment els límits de desenvolupament no estan establerts.

A més a més, tenen la tensió disminuïda, especialment en els músculs.

Trets físics:

Tenen una estatura mitjana més petita que la població general i certa tendència a l'obesitat.

Una de les característiques és la formació dels ossos del rostres, tenen l'occipital aplanat, tendència ha tenir els maxil·lars i la boca reduïda, també tenen una oclusió dental dolenta; orelles mal formades i sobresortints, pont nasal aplanat, també acostumen a tenir taques a l'iris (anomenades taques de Brushfield), i l'espai entre els marges de les parpelles de forma obliqua.

Pel que fa al tòrax, podem dir que tenen una cardiopatia congènita, principalment per defectes septals però sobretot de l'encoixinat endocràtic.

A la zona de l'abdomen i la pelvis, l'angle acetabular i ilíac disminuït i el penis petit.

Pel que fa a les mans i als peus, direm que tenen una palma singular, mans curtes i amples. Bucles cubitals en la majoria dels dits, en el 5é dit de la mà un únic doblec de flexió. En els dits dels peus, podem veure una petita distancia entre el dit I i II.

Altres signes observats amb freqüència significativa és que tenen una volta palatina alta, coll curt i ample, dents petites, llengua estriada, atrèsia intestinal (manca de continuïtat congènita de la llum de l'intestí) i anus imperforat.

Fisiologia :

Les persones amb SD tenen una elevada incidència en les cardiopaties congènites (del 35 al 50%), en general septacions incompletes i vàlvules defectuoses. També tenen més desequilibris hormonals i immunològics, problemes com per exemple el nivell d'hormones tiroides i una certa debilitat davant d'infeccions. També amb relativa freqüència problemes de naturalesa gastrointestinal i risc de leucèmia i malalties d'Alzheimer.

Controls mèdics

Segons l'edat de les persones amb SD s'han de dur uns determinats controls mèdics.

Nounat

En un nen acabat de néixer, el principal problema que pot tenir, és una cardiopatia congènita, és una de les malformacions més series dels nens amb SD. La freqüència de les cardiopaties congènites és del 40-50% dels nens. Les exploracions a què els han de sotmetre han d'incloure: un electrocardiograma, radiografies del tòrax i un ecocardiograma.

Un altre dels principals problemes és la malformació de l'aparell digestiu. La freqüència de casos és inferior a l'anterior (10%). També s'ha de tenir especial atenció en els òrgans dels sentits. Pel que fa a la visió, un 3% dels nens tenen cataractes congènites. Alguns nounats amb SD solen tenir pèrdua d'audició neurosensorial, per aquest motiu s'ha de realitzar exploracions durant els 6 primers mesos de vida.

Els primers problemes de la visió que poden aparèixer són: estrabisme, nistagmus, blefaroconjuntivitis i obstrucció del llagrimall.

També s'ha de fer exploracions otorrinobiològiques que s'han de realitzar de forma rutinària cada 6-12 mesos.

Infància

Els problemes de visió més freqüents durant la infància són: Errors de refracció (70%, miopia, hirmetropia i astigmatisme,,), l'estrabisme (35% dels casos), Nistagmus (5-15%) i Blefaroconjuntivitis de repetició (30%).

En aquesta edat poden aparèixer problemes de conducta o d'ansietat. A més a més d'aquests trastorns, n'existeixen dos més:

El trastorn d'hiperactivitat amb dèficit d'atenció, es dona a les persones amb SD amb més freqüència que a la població en general. Els símptomes característics són dificultat per concentrar-se, poca tolerància davant de frustracions, nivell d'activitat alt i inestabilitat emocional. Es tornen nens irritants, i difícils de tractar pels pares i professors. Aquest trastorn respon molt bé a fàrmacs estimulants.

I el trastorn tipus autista, també és més freqüent que en la població en general.

1.2. L'ull

L'ull és l'òrgan de visió, que rep les imatges de l'exterior i les transmet al cervell a través del nervi òptic i les vies òptiques. En el naixement, la grandària de l'ull és aproximadament tres quartes parts del d'un adult. El creixement és màxim durant el primer any i segueix creixent a un ritme ràpid però desaccelerat fins als tres anys. Després continua creixent fins a la pubertat però a un ritme més lent. El creixement posterior és insignificant. Les diferents parts de l'ull creixen a un ritme diferent. Les estructures de la part anterior de l'ull creixen proporcionalment menys que les de la part posterior.

L'ull, anomenat també globus ocular, és pràcticament esfèric d'uns 2,5 cm de diàmetre, estan subjectes i es poden moure gràcies als músculs extraoculars, que connecten la paret de la conca orbital amb la capa externa de l'ull.

Anatomia de l'ull

L'ull està protegit posteriorment per una cavitat del crani i, per la part anterior, per les estructures annexes: les celles, les pestanyes, les parpelles i els llagrimalls, que produeixen un líquid bactericida i lubricant. L'ull conté en el seu interior els líquids,

humor aquós i humor vitri, que li donen forma. L'ull està format per tres membranes (escleròtica, úvea i retina) i tres cambres.

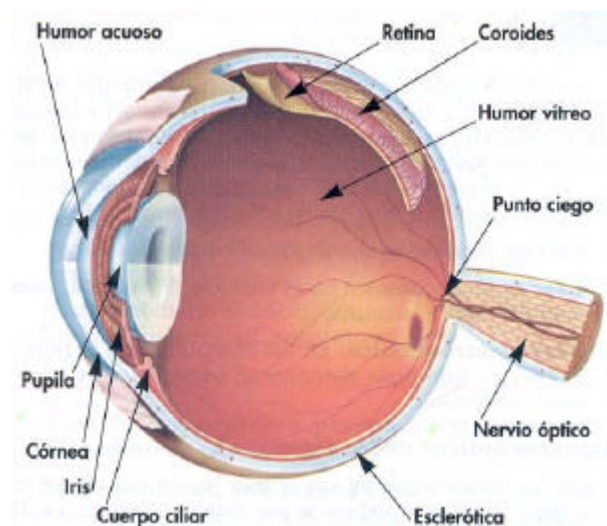
Escleròtica

És la capa més externa i dura. La seva funció és protectora i correspon a cinc sisenes parts de la superfície ocular. Aquesta membrana és fibrosa, blanca i opaca, excepte en la seva part anterior que és transparent (la còrnia), per permetre l'entrada de les imatges cap a la retina.

Úvea

És la capa del mig. És una membrana molt fina i fosca degut a la melanina. L'úvea és vascular, pigmentada i muscular, composta de diverses parts: la *coroide*, el *cos ciliar* i l'*iris*.

La coroide està en la part posterior i és una capa molt vascularitzada que revesteix les tres cinquenes parts posteriors del globus ocular. La seva funció és nodrir el cristal·lí i la retina, i fer de pantalla pigmentaria per evitar que la llum entri per on no ho ha de fer. Cap a la part anterior, la membrana es fa més gruixuda formant el cos ciliar.



El cos ciliar està format pels processos ciliars i els músculs ciliars. Els processos ciliars són els encarregats de la segregació del líquid que omple la part anterior i s'anomena humor aquós. Els músculs ciliars s'encarreguen de variar la curvatura del cristal·lí per enfocar a diferents distàncies, en aquests múscul se li uneixen lligaments per mantenir la posició del cristal·lí.

L'iris està situat a la zona anterior, darrera la còrnia, i té en el centre un orifici (o pupil·la). Té un color característic en cada persona i està format per fibres radials i circulars, la seva funció és regular la quantitat de llum que entra en l'interior de l'ull;

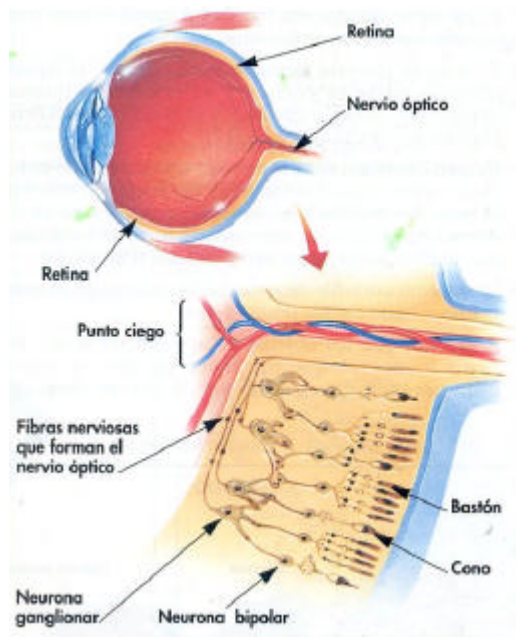
amb la contracció de les fibres permet l'obertura i el tancament de la pupil·la deixant passar més o menys llum.

Darrera l'iris es troba el cristal·lí, unit els músculs ciliars. És una lent d'enfocament i per tant la potència diòptrica per l'acció dels músculs ciliars.

Retina

La retina és la membrana més interna que recobreix la part posterior de l'ull on s'hauria de formar les imatges que veiem per poder veure-les amb nitidesa. A més a més, és complexa i està composta sobretot per cèl·lules nervioses. La retina està formada per dues capes, la pigmentada i la formada pels fotoreceptors.

La capa pigmentada és rica en melanina igual que la còrnia.



Les cèl·lules receptores sensibles a la llum es troben en la superfície exterior davant de la capa pigmentada. Hi han dos tipus de fotoreceptors: els cons (en tenim 7.000.000) i els bastons (125.000.000). Els cons estan agrupats al centre de la retina on té una petita taca groga anomenada **màcula**, dins d'aquesta es troba la **fòvea** zona on l'ull té l'agudes visual més gran, per tant, la capa sensorial de la fòvea està formada de cèl·lules en forma de con, al voltant d'aquestes n'hi ha en forma de bastons. A mesura que ens allunyem de l'àrea sensible es troben menys cèl·lules en forma de con. En les zones exteriors de la retina només trobem cèl·lules en forma de bastons.

El nervi òptic s'encarrega de travessar les múltiples capes que formen el globus ocular i d'enviar la informació visual des de l'ull al cervell. En la seva sortida, travessa les capes de l'ull. En el punt on travessa la retina no hi ha fotoreceptors i per tant, no hi ha capacitat per percebre estímuls lluminosos. Aquesta zona rep el nom de **punt cec**.

Les cambres que formen el globus ocular són tres:

L'**ANTERIOR** és la zona compresa entre la còrnia i l'iris. Està plena de l'humor aquós, és un líquid transparent produït pels processos ciliars i s'expulsen pels l'angle que formen l'iris i la còrnia.

La **POSTERIOR** també està plena de l'humor aquós, és la zona compresa entre l'iris i el cristal·lí on estan els processos ciliars.

I la **VÍTRIA** és la zona entre el cristal·lí i la retina, està plena d'un gel transparent anomenat humor vitri.

Funcionament de l'ull

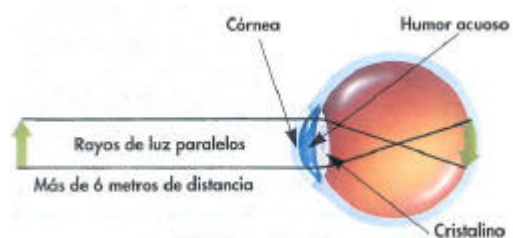
La visió consisteix en la captació d'energia lumínica visible per a les estructures oculars; l'enfoc sobre les estructures sensibles a la llum (retina) i la conversió pels receptors visuals en missatges nerviosos que, a través dels nervis òptics, arriben al cervell on són analitzats i interpretats.

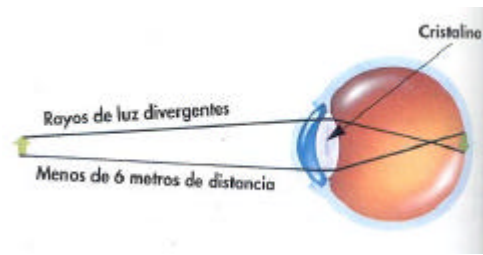
Podem diferenciar dues fases:

- a) **Fase òptica.** Consisteix en l'enfoc de les imatges visuals sobre la retina. La llum procedent d'un objecte s'ha d'enfocar a la retina, per percebre la seva imatge. Com que la distància entre l'objecte i l'ull és més gran que la distància entre la superfície de l'ull i la retina, els rajos de llum s'han de refractar per a què incideixi directament sobre la retina.

L'ull té dos sistemes per aconseguir enfocar els rajos de llum:

El sistema fix: està format per la còrnia, l'humor aquós, el cristal·lí i l'humor vitri. Aquest sistema és suficient pels objectes que es troben a una certa distància ja que els rajos de llum són paral·lels.





El sistema variables: l'enfocament està format pel cristal·lí. A mesura que l'objecte s'apropa a l'ull, els rajos es fan cada cop més divergents, pel que és necessari

augmentar la curvatura (més refracció) per que siguin enfocats a la retina. La modificació de la refracció és degut a la forma del cristal·lí, que pot ser més o menys esfèric, gràcies a la contracció i relaxació dels músculs ciliars.

La modificació de la refracció permet veure bé els objectes entre 6 metres (punt llunyà) i 15 centímetres (punt proper). A més de 6 metres no fa falta modificar la refracció i a menys de 15 centímetres ni modificant la refracció s'enfocarien bé.

- b) **Fase electroquímica.** En els receptors de la retina, l'energia lluminosa (fotons) es transforma en missatges nerviosos. Els cons i els bastons de la retina contenen uns pigments que al ser estimulats pels fotons, generen una energia que activa les neurones de la retina.

Els pigments dels cons són menys sensibles, és a dir, necessiten més llum per activar-se. Els bastons es poden activar quan no hi ha gairebé llum, i el fet que siguin més nombrosos que els cons afavoreix més la seva sensibilitat en els llocs foscos.

Els bastons només tenen un tipus de pigment, mentre que els cons en tenen tres, cadascun d'ells sensible a diferents longituds d'ones electromagnètiques. Aquestes longituds coincideixen amb els colors bàsics: verd, vermell i blau. Els bastons són més sensibles però només detecten imatges en tons grisosos.

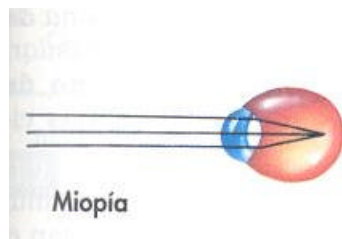
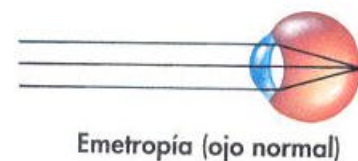
Quan canviem d'un lloc lluminós a un de fosc, o viceversa, es triga una estona fins que veiem alguna cosa i bastant de temps fins que ho veiem bé. Això és degut a dos mecanismes. El primer mecanisme és ràpid i és degut al tancament i l'obertura del iris. Amb això regulem l'entrada de llum a la retina. I el segon mecanisme és més lent i es deu a una acumulació o pèrdua dels pigments dels fotoreceptors.

Anomalies de l'ull

a) Defectes en la refracció i l'acomodació

Si s'enfoquen rajos lluminosos paral·lels sobre la retina amb l'ull en estat de repòs (no acomodat) i es forma correctament la imatge a la retina, existeix *emmetropia*. No és estrany, però és més normal l'estat oposat, *ametropia*, nom genèric de totes les anomalies de la refracció de l'ull per les quals les imatges no es formen a la retina.

Els defectes més freqüents són la miopia, la hipermetropia, l'astigmatisme i la presbícia. Aproximadament el 80% dels nens són fisiològicament hiperòptics en el naixement, el 5 % neixen miops.



MIOPIA: La longitud del globus ocular produeix que el focus de visió es realitzi per davant de la retina, sobretot en la visió d'objectes llunyans, per aquest motiu es produeix una visió borrosa. En canvi, els

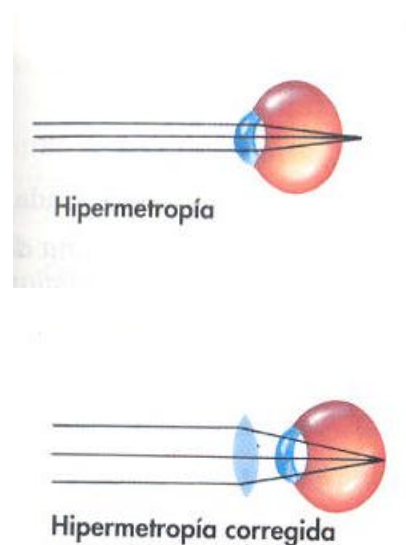
objectes propers no té problema per aquest motiu vulgarment s'anomenen "curts de vista". La paraula miopia prové del grec que significa ulls tancats perquè les persones amb miopia sovint tanquen lleugerament els ulls per veure millor de lluny.

La miopia pot ser deguda perquè el diàmetre anteroposterior de l'ull és relativament més llarg o perquè el poder de refracció de la còrnia o cristal·lí és molt més gran que el normal, o bé perquè el cristal·lí està dislocat cap endavant. Generalment l'ull miop és més gran del normal o la còrnia amb més curvatura del normal.

Els nens miops tendeixen a mantenir a prop els objectes i el material de lectura, prefereixen seure a prop de la pissarra i poden mostrar falta d'interès per activitats distants. És comú entre ells el frunziment de celles i l'estrabisme, perquè millora l'agudesia visual quan es redueix l'obertura de les parpelles. Es prescriuen lents còncaves (ulleres o lents de contacte) de potència apropiada per proporcionar visió clara i comoditat. S'aconsella la revisió anual. Habitualment, la miopia tendeix a augmentar en l'adolescència. Generalment es desenvolupa en l'època escolar i cap als 20 anys se sol estabilitzar. Fins llavors, es poden produir molts canvis i necessitar canvis d'ulleres o lents. Afecta tan als homes com a les dones i a més a més existeix una predisposició familiar. Els símptomes són visió borrosa dels objectes llunyans, mirar guerro, mal de cap i tensió ocular. El punt llunyà de visió clara varia inversament proporcional al grau de miopia; al augmentar la miopia s'apropa el punt distant de visió clara. Amb una miopia d'una diòptria, per exemple el punt llunyà d'enfocament clar és a 1 metre de l'ull, amb tres diòptries el punt llunyà es troba a només un terç de metre de l'ull.



HIPERMETROPIA: Si s'enfoquen rajos lluminosos per darrera de la retina amb l'ull en estat de repòs, existeix hipermetropia (o heperopia). El globus ocular és excessivament curt, o perquè la potència de refracció de la còrnia o del cristal·lí és menor que la normal, o perquè el cristal·lí està dislocat posteriorment. Generalment la grandària de l'ull és més petita del normal i la curvatura de la còrnia més plana.



En la hipermetropia s'utilitza l'acomodació per enfocar objectes situats a distància. Si l'esforç d'acomodació no és molt gran el nen tindrà visió clara i se sentirà còmode en la feina de visió propera i a distància. En els altres graus d'hipermetropia l'esforç d'acomodació és més gran, la visió pot ser borrosa i el nen pot queixar-se d'esforç ocular, cefalàlgies o fatiga. Altres manifestacions freqüents són l'estrabisme, fregament dels ulls, inflamació de les parpelles i desinterès per la lectura. Quan estan indicades es prescriuen lents convexes de potència suficient per a proporcionar una visió clara i còmode.

ASTIGMATISME: En aquest cas, la curvatura de la còrnia és irregular, el que produeix una visió distorsionada o de varis focus simultàniament. La imatge, que dona d'un punt enfocat, no és mai un punt sinó una taca. L'ull pot ser més



gran, més petit o de mida normal, el problema és que la còrnia no té forma esfèrica sinó una forma ovalada. Normalment, va acompanyada de miopia o hipermetropia. Hi ha una diferència en el

poder de refracció als diferents meridians de l'ull. La majoria dels casos és degut a una irregularitat en la curvatura de la còrnia, algun és degut a canvis en el cristal·lí. Un lleuger grau és comú i pot no produir símptomes. Amb graus més grans pot haver-hi deformació de la visió. Per poder tractar d'obtenir una imatge més clara, la persona utilitzarà l'acomodació, o el frunziment de celles, o l'estrabisme per obtenir un efecte d'orifici més petit, com en la miopia. Els símptomes són esforç ocular, cefalàlgia o fatiga, fregament dels ulls o hiperèmia de les parpelles, desinterès pel treball escolar i mantenir a prop el material de lectura, són manifestacions comuns a la infància. S'utilitzen lents cilíndriques o esferocilíndriques per proporcionar correcció òptica quan està indicat. Les lents poden ser necessàries constantment o només part del temps, segons el grau d'astigmatisme o la gravetat dels símptomes.

PRESBÍCIA: Amb l'edat es van perdent les capacitats d'enfocament de les lents oculars, se solen necessitar ulleres per llegir. Aquesta condició s'anomena presbícia o vista cansada. Aquest problema afecta a les persones entre 40 i 60 anys, quan el cristal·lí perd part de la seva flexibilitat i no pot focalitzar clarament els objectes propers. La presbícia es corregeix amb ulleres o lents de contacte.

ANISOMETROPIA: Existeix quan l'estat de refracció d'un ull és significativament diferent a l'estat de refracció de l'altre. El nen, generalment, utilitza l'ull normal o més normal de preferència a l'ull amb major error de refracció. Si no es corregeix pot conduir a ambliopia (o ull gandul) degut a la falta d'utilització. Per trobar el desenvolupament normal de l'ull fa falta la detecció i correcció el més aviat possible.

PARALISIS DE L'ACOMODACIÓ: La causa més freqüent és l'ús intencionat o inadvertit de substàncies cicloplègiques o tòpiques; s'inclou en elles tots els fàrmacs anticolinèrgics i els verins; així com les plantes i substàncies vegetals que contenen alcaloides naturals. Les causes neurògenes són les lesions que afecten al nervi oculomotor (el tercer nervi). El diagnòstic diferencial comprèn tumors, malalties degeneratives, lesions vasculars i traumatismes. Les malalties infeccioses poden afectar també a l'acomodació, per exemple, la diftèria pot causar paràlisi del múscul ciliar. Un defecte evident pot ser d'origen psicogen quan el nen fingeix incapacitat per llegir quan pot demostrar-se que la seva agudesia visual i la seva capacitat per enfocar són normals.

b) Trastorns del moviment i l'alineació de l'ull

En l'home, el desenvolupament de la visió normal en cada ull, el manteniment de l'alineació apropiada als eixos visuals (ortofòria) i la capacitat per integrar imatges dels dos ulls en una sola percepció visual que proporciona una altra percepció de la profunditat o estereòpsia, són íntimament interdependents.

Qualsevol variació sensitivomotora normal al començament de la vida pot donar resultats permanents de visió defectuosa o alineació ocular anormal. És important la ràpida detecció i tractament de l'estrabisme en nens i la valoració i tractament apropiat que requereixen el coneixement dels diferents tipus clínics d'estrabisme, els mètodes de detecció i els principis del tractament. Els dos tipus principals de desviació o una dolenta alineació de l'ull són heterofòria i heterotròpia.

La **HETEROFÒRIA** és una tendència latent a la mala alineació, l'ull només es desvia en certes dissociatives que dificulten el manteniment de la fusió normal. Les fòries són comuns i poden originar o no símptomes molestos com diplòpia transitòria, asteròpia (esforç muscular) o cefalàlgies. Quan el control de la desviació excedeix de l'amplitud de la fusió, de forma que es manifesta la desviació, la mala alineació s'anomena HETEROTRÒPIA (estrabisme) o tròpia. La alteració pot ser monocular o alternativa segons el tipus de visió i fixació. En l'estrabisme alternatiu es pot utilitzar un o l'altre ull per fixar o per la visió definitiva i l'altre es desvia, com que s'utilitza cada ull per torna, la visió es desenvolupa més o menys en cada ull. El pacient aprèn a suprimir la imatge en l'ull desviat, i fixa amb l'altre. Quan només s'utilitza (o es prefereix) un ull per la fixació i l'altre es desvia constantment, es coneix com estrabisme monocular o estrabisme dret o esquerre. El nen és propens a la ambliopia o visió central defectuosa de l'ull desviat, com a resultat del desús o mal ús.

L'estrabisme es descriu també segons la direcció de la desviació, la desviació convergent, creuament o rotació dels ulls cap a dins, es designa per prefix **eso** (esotròpia, esofòria) mentre que una desviació divergent o rotació dels ulls cap a fora amb el prefix **exo**. Les desviacions verticals s'indiquen amb el prefix **hiper** i **hipo**. Es poden mostrar aïllades o combinades, poden haver-hi desviacions torsionals o cicloverticals.

La classificació etiològica és complexa, degut a un coneixement incomplet dels factors o mecanismes causals. Es distingeixen paralític (no concomitant), no paralític (concomitant), acomodatiu i no acomodatiu.

L'estrabisme paralític és degut a la debilitat o paràlisi d'un o més músculs extraoculars, que empitjora la desviació amb la mirada fixa en el camp d'acció del múscul afectat. Per tant, en el cas d'una parèsia abducent dreta els ulls apareixen creuats al mirar a la dreta, però rectes (ortofòrics) al mirar a l'esquerra. La manifestació subjectiva és diplopia (visió doble), per evitar aquesta molesta sensació pot girar el cap per evitar mirar en la direcció del múscul parètic, pot tapar-se un ull per eliminar la doble imatge. La paràlisi dels músculs extraoculars són signes d'un greu procés patològic, l'aparició d'un estrabisme no concomitant pot ser el primer signe d'un tumor intracraneal, un procés infecciós o parainfecciós (meningitis, encefalitis, neuritis), una malaltia desmielinitzant o neurodegenerativa, miastènia greu o miopatia progressiva. L'estrabisme paralític congènit és degut a defectes del desenvolupament dels nuclis o fibres de nervis cranials, anomalies musculars, síndromes d'infecció congènita o traumatisme al néixer.

L'estrabisme no paralític és el més comú. No hi ha cap defecte en l'acció dels músculs i el grau de desviació és constant o relativament constant en totes les direccions de la mirada. La majoria de les esotròpies congènites o infantils són del tipus no paralític o concomitant, aquest tipus es tracta millor quirúrgicament, però el tractament efectiu ha d'abastar el de qualsevol ambliopia. Alguns casos d'estrabisme no paralític és degut a defectes oculars o visuals adjacents, com són les cataractes, lesions del nervi òptic o de la màcula, grans errors de refracció o errors de refracció asimètrics (anisometria). Quan sigui possible s'ha de corregir primer l'afecció ocular adjacent (aspiració de les cataractes, correcció de la refracció,...) i en casos seleccionats cirurgia cosmètica per col·locar els ulls bé.

Un tipus d'estrabisme no paralític és l'estrofia acomodativa. Aquest tipus depèn de la relació entre els reflexos d'acomodació origina sobreconvergència o creuament d'ulls; en alguns casos també hi ha un trastorn de la relació de curta a llarga distància, el grau de creuament a curta distància és major que el llarg. Aquest tipus apareix a l'edat de 2 a 3 anys, amb una variació de començament aproximadament de 6 mesos a 7 o 8 anys. La majoria de nens

afectats tenen un cert grau de hiperopia (presbiopia); en molts casos, l'estrabisme pot ser controlat amb lents que convergeixen la hiperopia, alguns necessiten lents bifocals.

c) Altres problemes de visió

Altres problemes de visió que pateixen els nens amb SD és el Nistagmus i la blefaroconjuntivitis de repetició.

NISTAGMUS: És un moviment incontrolat i involuntari de l'ull. Normalment afecta als dos ulls i es posa de manifest al fixar la vista en una determinada direcció. A vegades el control del cervell sobre els moviments de l'ull és pobre, resultant la impossibilitat de mirar fixament un objecte. Algunes formes de nistagmus estan associades a una visió reduïda, com pot passar amb els albins, o en aquells casos amb cicatrius a la retina o patologies del nervi òptic. Rarament pot aparèixer com a resultat de tumors cerebrals o alteracions neurològiques greus. Existeixen dues formes de nistagmus: el motor i el sensorial. El nistagmus motor comença entre les 6 setmanes i els 3 mesos d'edat. No hi ha antecedents a la família de moviments oculars similars als del pacient. El moviment és usualment horitzontal. A vegades la convergència o la mirada en una direcció redueix la intensitat del nistagmus i millora la visió. Afortunadament el pacient no veu els objectes del seu entorn movent-se tal i com ho fan els seus ulls. L'agudesia visual pot estar reduïda en la visió llunyana però gairebé normal en la propera. El nistagmus sensorial està associat a una disminució de la visió sense cap causa aparent. Normalment comença a les 6 i 8 setmanes d'edat. Els ulls semblen que vaguegin, a vegades lentament i de vegades ràpidament. Alguns cops els ulls marxen cap amunt. La disminució de visió acostuma a ser tractable i el pacient pot arribar a portar una vida perfectament normal.

BLEFAROCONJUNTIVITIS: És una inflamació de les parpelles i la conjuntiva. Habitualment és d'origen infeccios o al·lèrgic.

2. Objectius

El projecte tracta d'il·lustrar la metodologia dels models lineals mixtos per l'anàlisi d'experiments de mesures repetides i longitudinals. L'anàlisi d'aquests experiments suposa l'existència de dependència entre les observacions d'un mateix individu.

Aquest treball està format per una part teòrica on s'explica la metodologia dels models lineals mixtos i per una part pràctica que reflecteix la metodologia d'anàlisi. Per dur a terme la part pràctica utilitzarem un conjunt de dades reals on aplicarem els conceptes explicats. A continuació, analitzarem aquestes dades mitjançant el paquet estadístic R i concretament la llibreria nlme.

Podríem utilitzar els models lineals clàssics (models lineals generalitzats), però aquest tipus d'anàlisi requereix que les dades segueixin una mostra aleatòria simple i no sempre és així. Moltes vegades, l'experiència ens diu que les dades estan correlacionades a diferents nivells.

Al llarg de la carrera ens han explicat diferents maneres d'analitzar les mesures repetides al llarg del temps, com són els models **ANOVA per a mesures repetides** i els models **MANOVA**. Aquest models utilitzen el mètode dels moments per estimar els paràmetres.

ANOVA per a mesures repetides

En aquest model, les mesures repetides actuen com a bloc, d'aquesta manera es podem utilitzar com a factors en un anàlisi de la variància clàssic. Aquests models requereixen menys subjectes que un disseny completament aleatori i permeten eliminar la variància residual deguda a les diferències entre els subjectes. Les mesures repetides per un individu formen un bloc i, per tant, no es té en compte les interaccions entre el factor bloc i els altres efectes d'interès.

MANOVA

Aquest model és una extensió del model ANOVA per als casos en que la variable resposta és un vector, per tant, tot l'anàlisi del MANOVA es fa de forma vectorial.

Un dels principals inconvenients és que, tant el cas de l'ANOVA per a mesures repetides com el cas del MANOVA, requereixen dades balancejades i mesurades en el mateix instant de temps.

Els models lineals mixtos ens permeten modelitzar dades no balancejades i mesurades en diferents instants de temps. En la bibliografia, es poden trobar explicats sota diferents noms (**Model lineal mixt**, **Models de coeficients aleatoris** i **models a dues etapes**), però bàsicament tenen la mateixa filosofia de treball amb les dades.

Model lineal mixt (mixed model)

Aquest model ens permet treballar amb les dades de tal manera que la part aleatòria inclou informació de cada individu. Això ens permet treballar amb dades no balancejades.

La formulació és:

$$Y_i = \underbrace{X_i \mathbf{b}}_{\text{Part fixa}} + \underbrace{Z_i \mathbf{b}_i}_{\text{Part aleatòria}} + \mathbf{e}_i$$

Models de coeficients aleatoris (Random coefficient model)

En aquest model, cada individu comparteix una part de la informació poblacional a la vegada que cadascun d'ells té un coeficient aleatori. Així, aquest model admet la variabilitat deguda a cada individu.

En aquesta formulació, considerem que X_i i Z_i són iguals (notades com \tilde{X}_i), d'aquesta manera podem escriure el model com:

$$Y_i = \tilde{X}_i (\mathbf{b} + \mathbf{b}_i) + \mathbf{e}_i$$

Models a dos etapes o multinivell (two-stage o multilevel)

Aquest tipus de model s'acostuma a utilitzar en mesures biomèdiques i en l'àmbit dels estudis sociològics i educacionals. Habitualment es registren diverses mesures al llarg del temps sobre una mateixa unitat experimental.

En la formulació d'aquest model tenim una modelització a dos nivells, els coeficients i les observacions:

$$Y_i = X_i \mathbf{b}_i + \mathbf{e}_i$$
$$\mathbf{b}_i = \mathbf{b} + b_i$$

Els paràmetres del model són $\mathbf{b}, D, \mathbf{s}$:

$$\mathbf{e}_i \sim N(0, \mathbf{s}^2)$$
$$b_i \sim N(0, \mathbf{s}^2 D)$$

on \mathbf{b} descriu la part fixa i $D i \mathbf{s}$ descriuen la distribució dels efectes aleatoris i dels residus.

Durant la recerca de bibliografia en l'àmbit aplicat per aquest tipus de dades, s'han trobat referències on s'estudiaven dades que es modelitzaven mitjançant distribucions binomials sense mesures repetides.

Existeixen estudis amb dades longitudinals semblants a les dades utilitzades en el present treball. En l'article de *Mary Cregg, J. Margaret Woodhouse i altres*², es resumeix la informació d'un dels ulls en una nova variable. En un altre article, *Olav H Haugen, Gunnar Høvdning, Isa Lundström*³, no s'utilitza la informació d'un dels ulls, d'aquesta manera eviten les correlacions entre les dades d'un nivell en concret. En tots aquests casos, els protocols estaven ben definits per tal que les diferents mesures es prenguessin en el mateixos instants de temps.

Durant el temps destinat al procés de recerca bibliogràfica, no s'ha posat de manifest cap antecedent d'utilització de models lineals mixtos en aquest context.

Els models lineals mixtos ens permeten modelitzar dades longitudinals amb mesures repetides preses en diferents instants de temps, i d'aquesta manera permeten estudiar al màxim la informació recollida.

3. Materials i mètodes

L'objectiu d'aquest estudi és il·lustrar amb un exemple, un model d'anàlisi de dades basat en el model lineal mixt. Per aquest motiu en aquest apartat definirem els models estadístics aplicats, comentarem algunes referències històriques i explicarem el programa utilitzat. La terminologia utilitzada per explicar el models lineals mixtos és la que es mostra al llibre *Mixed-effect Models in S and S-Plus* (José C. Pinheiro i Douglas M. Bates, 2000)¹.

A la base de dades, tenim una mateixa variable registrada en diferents moments al llarg del temps, això es coneix com a dades longitudinals. En concret, tenim índexs de refracció per a cadascun dels individus en diferents dies. Les mesures per a cada individu estan correlacionades.

Ens interessa establir la dependència funcional entre l'edat i l'índex de refracció. La problemàtica més gran és que el nombre d'observacions per individu no és idèntic i les dades estan recollides en intervals de temps irregulars. És per aquest motiu que no podem aplicar l'anàlisi de la variància; ja que, tot i que pot contemplar mesures repetides, necessita que el nombre d'observacions per element sigui idèntic.

Per poder modelitzar amb un nombre de paràmetres i poder extraure conclusions, utilitzarem els models mixtos. Podríem ajustar un model per a cada individu però aconseguiríem un model sobreparametritzat. Si ajustéssim un model poblacional, no tindríem en compte la correlació entre les observacions d'un mateix individu.

Existeixen estudis similars que no apliquen els models lineals mixtos. En l'estudi de *Mary Cregg, J. Margaret Woodhouse i altres*² s'ometen les dades d'un dels ulls sense cap tipus de criteri específic, d'aquesta manera no tenen dades correlacionades. La variable resposta és la mitjana de l'índex refractari entre la primera visita i la última visita. En tots els nens, el nombre de mesos transcorreguts entre la primera i la última visita són els mateixos, 30 mesos. En l'estudi de *Olav H Haugen, Gunnar Høvdning, Isa Lundström*³ s'ometen les dades de l'ull esquerre i classifiquen els nens segons l'evolució de l'ull dret entre la primera i la última visita. Per tenir les dades mesurades en el mateix instant de temps, fixen el seguiment dels pacients en 30 mesos, sense

tenir en compte l'edat del nen en la primera visita. Les dades passen a ser una taula de contingència on es classifiquen els nens segons la tendència de l'índex refractari i ajusten un model utilitzant la chi-quadrat (χ^2). En aquests dos articles, utilitzen l'ordre de les visites com a dades longitudinals, quan podrien utilitzar la variable edat que, a part de ser una dada longitudinal, ens indica la cronologia de les mesures.

3.1. Formulació del model

Els models lineals multinivell són models estadístics de paràmetres que varien en més d'un nivell. Aquests models s'usen per estimar la variància dins del grup i entre els grups. En aplicacions psicològiques, els diferents nivells podrien ser les preguntes a un qüestionari, individus i família.

Considerem les mesures repetides com a dades amb múltiples observacions sobre la mateixa unitat experimental.

Els models multinivell amb mesures repetides tenen aplicacions amb **dades longitudinals**, direm que tenim dades longitudinals quan tenim moltes mesures d'un mateix individu que varien al llarg del temps, també poden variar sobre l'espai. Aquesta situació fa que les observacions sobre la mateixa unitat estiguin correlacionades. Un exemple seria els estudis de creixement per tal de veure canvis en els individus i les diferències entre ells.

Els models lineals mixtos estan formats pels **efectes fixes** i els **aleatoris**. Tant els efectes fixes com els aleatoris són lineals. A l'incorporar els efectes aleatoris al model, estem incorporant un error addicional, aquest error és degut a la correlació de les observacions dins del mateix grup. Assumim que els *efectes fixes* tenen un nombre finit de nivells i s'utilitzen tots en el model. En canvi, pels *efectes aleatoris*, assumim que tenim infinits nivells, i per tant, només utilitzem una mostra d'aquesta població. Els efectes aleatoris sorgeixen quan tenim més d'una unitat experimental, aleshores tindrem observacions correlacionades dins de la unitat o del nivell.

Altres aplicacions dels models d'efectes mixtos són els dissenys d'experiments jeràrquics, com els dissenys de parcel·les dividides (Split-plot). Algunes vegades el

interès principal és mesurar les variàncies de les variables aleatòries, per aquest motiu rep el nom de component de la variància (Searle et al., 1992).

L'aproximació dels models mixtos és coneguda com a modelització condicional; els coeficients de regressió s'apliquen a cada individu però no necessàriament a la població.

El cas més simple dels models lineals mixtos es pot considerar quan hi ha dues etapes de mostreig. Al primer estadi, les unitats es prenen aleatòriament dins d'una població i , en el segon, es fan diferents mesures a cada unitat.

Els models lineals mixtos han guanyat popularitat en l'última dècada, un mica perquè s'han desenvolupat softwares eficients per ajustar i analitzar aquests models. Pels models lineals i no lineals la llibreria nlme del programa R n'és un exemple.

Formulació del model simple (Laird and Ware, 1982)

El model lineal mixt simple s'expressa a partir de n_i dimensions. La variable resposta y_i per l' i -èsim grup és:

$$Y_i = X_i \mathbf{b} + Z_i \mathbf{b}_i + \mathbf{e}_i$$

on $\mathbf{b}_i \sim N(0, \mathbf{y})$ i $\mathbf{e}_i \sim N(0, \mathbf{s}^2 \cdot Id)$, $i=1, \dots, M$ i \mathbf{b} és el vector d'efectes fixes de p dimensions i \mathbf{b}_i és el vector dels efectes aleatoris de q dimensions. Els regressors són X_i pels efectes fixes amb (n_i, p) dimensions i Z_i pels efectes aleatoris (n_i, q) . El vector d'error dins del grup és \mathbf{e}_i (n_i dimensions) amb distribució esfèrica de Gauss (distribució normal multidimensional amb variància constant).

Suposem que la variància dels errors és constant ($Var(\mathbf{e}_i) = \mathbf{s}^2 Id$) i, que els efectes aleatoris i els errors són independents per diferents grups i per dins del mateix grup. A l'apartat sobre l'extensió dels models lineals mixtos veurem que la variància dels errors no ha de ser necessàriament constant.

Sempre que la distribució dels efectes aleatoris (b_i) segueix normal de mitjana 0, per tant, la matriu de variàncies i covariàncies és \mathbf{y} . Aquesta matriu és simètrica i definida semipositiva, és a dir, els valors propis no poden ser negatius.

Per treballar amb la matriu de variàncies i covariàncies és millor fer-ho amb un factor de precisió:

$$\frac{\mathbf{y}^{-1}}{1/s^2} = \Delta^T \Delta$$

Si la matriu \mathbf{y} és definida positiva, aleshores existeix Δ , però no és única. El factor de Cholesky $s^2 \mathbf{y}^{-1}$ és un valor possible per Δ , aquesta matriu rep el nom de factor relatiu de precisió, perquè és un factor de precisió de la matriu \mathbf{y}^{-1} dels efectes aleatoris relatius a $1/s^2$ dels errors \mathbf{e}_i .

Formulació del model Multinivell

La formulació del model simple es pot ampliar a multinivell, afegint nivells d'efectes aleatoris. En el cas de dos efectes aleatoris la variable resposta serà: y_{ij} on $i=1, \dots, M$ i $j=1, \dots, M_i$ on M són els nombres del primer nivell i M_i són els nombres de grups de segon nivell dins del primer nivell. La grandària de y_{ij} és de n_{ij} .

La matriu d'efectes fixes és X_{ij} ($i=1, \dots, M$; $j=1, \dots, M_i$) ($n_{ij} \times p$), el primer nivell és b_i (q_1) i el segon b_{ij} (q_2); les matrius són $Z_{i,j}$ i Z_{ij} .

$$Y_{ij} = X_{ij} \mathbf{b} + Z_{i,j} b_i + Z_{ij} b_{ij} + \mathbf{e}_{ij} \quad (i=1, \dots, M; j=1, \dots, M_i) \quad \text{on} \quad b_i \sim N(o, \mathbf{y}_1)$$

independents per i's diferents, $b_{ij} \sim N(o, \mathbf{y}_2)$ independent per i o j diferents a i pel nivell 1.

Si tenim Q nivells d'efectes aleatoris la formulació és:

$$Y_{ijk} = X_{ijk} \mathbf{b} + Z_{i,jk} b_i + Z_{ij,k} b_{ij} + Z_{ijk} b_{ijk} + \mathbf{e}_{ijk}$$

amb $i = 1, \dots, M$; $j = 1, \dots, M_i$; $k = 1, \dots, M_{ij}$.

on $b_i \sim N(o, \mathbf{y}_1)$, $b_{ij} \sim N(o, \mathbf{y}_2)$, $b_{ijk} \sim N(o, \mathbf{y}_3)$ i $\mathbf{e}_{ijk} \sim N(0, \mathbf{s}^2 \mathbf{I})$.

Igual que al nivell simple podem expressar \mathbf{y}_q ; $q = 1, \dots, Q$ a termes relatius Δ_q .

3.2. Estimació dels paràmetres

Molts procediments estadístics suposen que les dades segueixen algun model matemàtic que es defineix mitjançant una equació, de la que es desconeix algun dels seus paràmetres que es calculen o estimen a partir de la informació obtinguda en un estudi dissenyat per aquesta finalitat. Existeixen diferents procediments per estimar els coeficients d'un model de regressió, o per estimar els paràmetres d'una distribució de probabilitat. De tots aquests procediments, el més utilitzat per la seva versatilitat és el de màxima versemblança (Maximum Likelihood), ja que es pot aplicar a una gran quantitat de situacions.

➤ Màxim Likelihood

Suposem que tenim una mostra aleatòria simple de X on X segueix una distribució F de paràmetres \mathbf{J} . Aleshores l'estimador de màxima versemblança és aquell valor que minimitza el pes de la probabilitat a la hora d'escollir la mostra.

En el nostre cas, la funció de versemblança és la probabilitat de la funció de densitat donats els paràmetres:

$$L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2 / y) = p(y / \mathbf{b}, \mathbf{J}, \mathbf{s}^2)$$

on L és la funció likelihood

y és la variable resposta (vector n -dimensional)

$$N = \sum_{i=1}^M n_i$$

El estimador de màxima versemblança és aquell que maximitza la funció de versemblança.

La funció del likelihood és:

$$L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2 / \mathbf{y}) = \prod_{i=1}^M p(y_i / \mathbf{b}, \mathbf{J}, \mathbf{s}^2)$$

Com que la funció és monòtona creixent i podem suposar que és positiva, el nostre estimador serà un estimador de màxima versemblança si maximitza la funció de log-Likelihood. Aquest càlcul, inclou el càlcul de la marginal de les dades respecte els efectes aleatoris que no són observables.

$$\ell(\mathbf{b}) = \log(L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2 / \mathbf{y}))$$

Donat \mathbf{J} , els valors \mathbf{b} i \mathbf{s}^2 (màx. Likelihood) són:

$$\begin{aligned} \hat{\mathbf{b}}(\mathbf{J}) &= \left(\sum_{i=1}^M \mathbf{X}_i^T \Sigma_i^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^M \mathbf{X}_i^T \Sigma_i^{-1} y_i \\ \mathbf{s}^2(\mathbf{J}) &= \frac{\sum_{i=1}^M (y_i - \mathbf{X}_i \hat{\mathbf{b}}(\mathbf{J}))^T \Sigma_i^{-1} (y_i - \mathbf{X}_i \hat{\mathbf{b}}(\mathbf{J}))}{N} \end{aligned}$$

El mètode de màxima versemblança obté estimadors que són funcions d'estadístics suficients sempre que l'estimador obtingut sigui únic.

Com que estimem la part fixa com a part de la mitjana i de la part aleatòria simultàniament, obtenim unes estimacions de la part aleatòria i de la part residual que estan esbiaixades. Si convergissin, la mitjana no seria esbiaixada.

Per evitar aquest problema, molts analistes prefereixen utilitzar el **Restricted maximum likelihood (REML)** quan estimen els components de la variància (\mathbf{s}^2 i \mathbf{s}_b^2), ja que amb el **màxim Likelihood** s'acostuma a sobreestimar la variància.

➤ **Restricted Likelihood**

Existeixen moltes vies per definir el REML. Una manera és la de Laird and Ware (1982):

$$L_R(\mathbf{J}, \mathbf{s}^2/y) = \int L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2/y) d\mathbf{b}$$

Integrem dins d'un context Bayesià, això significa que assumim a priori una distribució local uniforme pels efectes fixes (\mathbf{b}) i els integrem fora de la versemblança.

Si utilitzem les mateixes tècniques que en el cas del mètode de màxima versemblança i apliquem logaritmes, obtenim una distribució de la qual podem obtenir l'estimació de restricted maximum Likelihood:

$$l_R(\mathbf{J}/y) = l_R(\mathbf{J}, \mathbf{s}_R^2(\mathbf{J})/y)$$

Els estimadors obtinguts pel mètode del restricted maximum Likelihood són estimadors sense biaix.

3.3. Inferència

Abans de començar a fer inferència hem de determinar que les distribucions que hem assumit siguin vàlides per a les nostres dades. En els models mixtos assumim dues distribucions:

- a. Els errors són independents, segueixen una distribució normal idènticament distribuïda, amb mitjana 0 i variància σ^2 i, són independents dels efectes aleatoris.
- b. Els efectes aleatoris tenen una distribució normal amb mitjana 0 i matriu de covariàncies \mathbf{y} i són independents per grups diferents.

Per determinar la significació dels paràmetres i comprovar si podem simplificar el model, utilitzem els tests d'hipòtesis i els intervals de confiança.

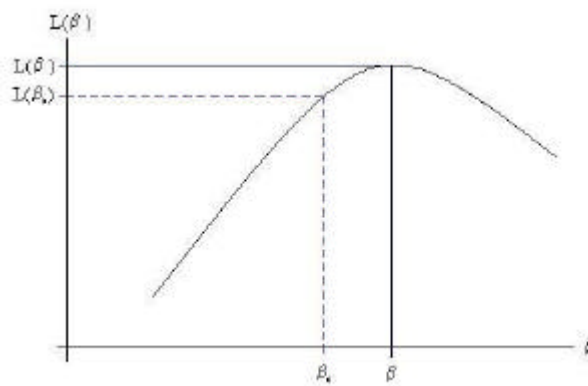
Test d'hipòtesis

Utilitzarem dos tests d'hipòtesis, **Likelihood Ratio** i **test de Wald**.

1.- Likelihood Ratio Test (Lehman, 1986)

Aquest test s'utilitza amb els models realitzats amb REML per comparar dos models. L'estimació dels efectes fixes ha de ser la mateixa.

Suposem que L_2 és el model més general i L_1 és el model amb Likelihood més restrictiu (menys termes), tenim que $L_2 \supset L_1 \Rightarrow \log L_2 \supset \log L_1$.



El test likelihood ratio és:

$$2 \log \left(\frac{L_2}{L_1} \right) = 2 [\log(L_2) - \log(L_1)] > 0$$

Si K_i és el nombre de paràmetres estimats en el model i , aleshores la distribució asimptòtica de l'estadístic de Likelihood Ratio Test (LRT), sota la hipòtesis nul·la (model més restrictiu adequat), segueix una χ^2 ($K_2 - K_1$ graus de llibertat):

$$LRT \sim \chi^2_{K_2 - K_1}$$

Utilitzem el Ratio Test per avaluar la significació dels efectes aleatoris.

2.- Test de Wald

Es vol realitzar un contrast d'hipòtesi respecte un paràmetre \mathbf{J} , pel que existeix un estimador màxim versemblant $\hat{\mathbf{J}}$:

$$\begin{aligned} H_0 : \mathbf{J} &= \mathbf{J}_0 \\ H_1 : \mathbf{J} &\neq \mathbf{J}_0 \end{aligned}$$

El contrast es pot realitzar a partir de la distància entre \mathbf{J}_0 i l'estimador màxim versemblant $\hat{\mathbf{J}}_n$. Sota H_0 la distribució aproximada de $\hat{\mathbf{J}}_n$ és:

$$\hat{\mathbf{J}}_n \approx N_k(\mathbf{J}_0, (\mathbf{I}_n(\mathbf{J}_0))^{-1})$$

i l'estadístic de Wald mesura la distància de Mahalanobish entre \mathbf{J}_0 i $\hat{\mathbf{J}}_n$.

$$W_n = (\hat{\mathbf{q}}_n - \mathbf{J}_0)^t \mathbf{I}_n(\mathbf{J}_0) (\hat{\mathbf{q}}_n - \mathbf{J}_0)$$

La matriu $\mathbf{I}(\mathbf{J}_n)$ es pot substituir per un estimador consistent de la matriu de informació. Per exemple, es pot utilitzar $\mathbf{I}(\hat{\mathbf{J}}_n)$ o la matriu de informació observada \mathbf{O}_n .

$$\mathbf{I}(\mathbf{J}_n) = E \left(- \frac{\partial^2 \ell(\mathbf{J}, x)}{\partial \mathbf{J}^2} \right)$$

L'estadístic del test de Wald sota H_0 segueix una chi-quadrat ($W_n \approx \chi^2_{k,a}$).

Rebutgem la hipòtesis nul·la si $W_n(X_n) > \chi^2_{k,a}$

Aquest test s'utilitza normalment per verificar els efectes fixes, encara que també pot servir pels efectes aleatoris però no sempre aconseguir convergir.

Intervals de confiança

Anomenem interval de confiança a un interval de valors al voltant d'un paràmetre mostral, amb un nivell de confiança determinat, on se situarà el paràmetre poblacional a estimar. Si α és l'error aleatori que es vol permetre, la probabilitat serà $1-\alpha$.

Si $L(X_n, \mathbf{J})$ és la funció de versemblança d'una mostra de grandària n de

$X \sim f(x, \mathbf{J})$, $\hat{\mathbf{J}}_n$ és l'estimador màxim versemblant de \mathbf{J} i la informació de Fisher

$I_n(\mathbf{J}) = -E_{\mathbf{J}} \left(\frac{\partial^2}{\partial \mathbf{J}^2} \log L(\mathbf{J}, X_n) \right)$ que sobre \mathbf{J} conte una mostra n , aleshores:

$$\frac{(\hat{\mathbf{J}}_n - \mathbf{J})}{\sqrt{(I_n(\mathbf{J}))^{-1}}} \xrightarrow{D} Z \sim N(0,1)$$

El interval de confiança aproximadament de $1-\alpha$ per \mathbf{J} és:

$$\hat{\mathbf{J}}_n \pm z_{\alpha/2} \sqrt{(\hat{I}_n(\mathbf{J}))^{-1}}$$

on $\hat{I}_n(\mathbf{J})$ és un estimador consistent de la informació de Fisher.

Prediccions

Les prediccions varien segons el nombre de nivells.

Nivell $\geq K$:

$$E[y_n] = X_h^T \mathbf{b}$$

Nivell 1:

$$E[y_h(i)/b_i] = X_h^T \mathbf{b} + Z_h(i)^T b_i$$

on $Z_h(i) \rightarrow$ Vector covariàncies dels efectes aleatoris.

Nivell 2:

$$E[y_h(i)/b_i b_{ij}] = X_h^T \mathbf{b} + Z_h(i)^T b_i + Z_h(i, j)^T b_{ij}$$

on $Z_h(i, j) \rightarrow$ Covariàncies del nivell j en el nivell i .

Els millors estimadors sense biaix per fer aquestes prediccions són:

$$\hat{y}_h = X_h^T \hat{\mathbf{b}}(\mathbf{J})$$

$$\hat{y}_h(i) = X_h^T \hat{\mathbf{b}}(\mathbf{J}) + Z_h(i)^T \hat{b}_i(\mathbf{J})$$

$$\hat{y}_h(i, j) = X_h^T \hat{\mathbf{b}}(\mathbf{J}) + Z_h(i)^T \hat{b}_i(\mathbf{J}) + Z_h(i, j)^T \hat{b}_{ij}(\mathbf{J})$$

A la pràctica, \mathbf{J} és desconeguda i la substituïm pel seu estimador màxim versemblant o per l'estimador REML.

3.4. Aspectes tècnics (computacionals)

Computacionalment, existeixen diferents mètodes per estimar els paràmetres, entre ells el **Màxim Likelihood** i el **Restricted Likelihood**.

➤ Màxim Likelihood

En el model simple (un únic nivell d'objectes aleatoris), els paràmetres a estimar són: \mathbf{b}, \mathbf{s}^2 i \mathbf{J} és el paràmetre que utilitzem per determinar Δ .

La funció del Likelihood és la probabilitat de la funció de densitat donats els paràmetres:

$$L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2 / y) = p(y / \mathbf{b}, \mathbf{J}, \mathbf{s}^2)$$

on L és la funció likelihood

y és la variable resposta (vector n-dimensional)

$$N = \sum_{i=1}^M n_i$$

Els efectes aleatoris no observables b_i formen part del model, hem d'integrar la densitat condicional donats els efectes aleatoris, així trobem la densitat marginal de les dades.

$$\begin{aligned} L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2 / y) &= \prod_{i=1}^M p(y_i / \mathbf{b}, \mathbf{J}, \mathbf{s}^2) = \\ &= \prod_{i=1}^M \int p(y_i / b_i, \mathbf{b}, \mathbf{s}^2) \cdot p(b_i / \mathbf{J}, \mathbf{s}^2) db_i \end{aligned} \quad (1)$$

on la densitat condicionada de y és normal multivariant:

$$p(y_i / b_i, \mathbf{b}, \mathbf{s}^2) = \frac{\exp\left(-\|y_i - X_i \mathbf{b} - Z_i b_i\|^2 / 2\mathbf{s}^2\right)}{(2\pi\mathbf{s}^2)^{n_i/2}} \quad (*)$$

i la densitat marginal de b_i és també normal multivariant

$$p(b_i/J, s^2) = \frac{\exp(-b_i^T Y^{-1} b_i / 2)}{(2p)^{q/2} \sqrt{|Y|}} = \frac{\exp(-\|\Delta b_i\|^2 / 2s^2)}{(2ps^2)^{q/2} abs|\Delta|^{-1}} \quad (**)$$

on $|\Delta|$ denota el determinant de la matriu Δ .

Substituïm la densitat condicionada de y (*) i la densitat marginal de b_i (**) a la densitat marginal de les dades **(1)**.

$$\begin{aligned} L(\mathbf{b}, \mathbf{J}, s^2 / y) &= \\ &= \prod_{i=1}^M \frac{abs|\Delta|}{(2ps^2)^{n_i/2}} \int \frac{\exp(-\|y_i - X_i \mathbf{b} - Z_i b_i\|^2 + \|\Delta b_i\|^2 / 2s^2)}{(2ps^2)^{q/2}} db_i = \\ &= \prod_{i=1}^M \frac{abs|\Delta|}{(2ps^2)^{n_i/2}} \int \frac{\exp(-\|\tilde{y}_i - \tilde{X}_i \mathbf{b} - \tilde{Z}_i b_i\|^2 / 2s^2)}{(2ps^2)^{q/2}} db_i \end{aligned} \quad (2)$$

on $\tilde{y}_i = \begin{bmatrix} y_i \\ 0 \end{bmatrix}$, $\tilde{X}_i = \begin{bmatrix} X_i \\ 0 \end{bmatrix}$, $\tilde{Z}_i = \begin{bmatrix} Z_i \\ \Delta \end{bmatrix}$ són vectors i matrius augmentats del

model.

Si ens fixem en l'equació anterior, l'exponent de l'integral té forma de suma de quadrats residuals, podem determinar \hat{b}_i minimitzant aquesta suma. La solució és:

$$\hat{b}_i = (\tilde{Z}_i^T \tilde{Z}_i)^{-1} \tilde{Z}_i^T (\tilde{y}_i - \tilde{X}_i \mathbf{b})$$

La norma al quadrat s'expressa:

$$\begin{aligned} \|\tilde{y}_i - \tilde{X}_i \mathbf{b} - \tilde{Z}_i \hat{b}_i\|^2 &= \|\tilde{y}_i - \tilde{X}_i \mathbf{b} - \tilde{Z}_i \hat{b}_i\|^2 + \|\tilde{Z}_i (b_i - \hat{b}_i)\|^2 = \\ &= \|\tilde{y}_i - \tilde{X}_i \mathbf{b} - \tilde{Z}_i \hat{b}_i\|^2 + (b_i - \hat{b}_i)^T \tilde{Z}_i^T \tilde{Z}_i (b_i - \hat{b}_i) \end{aligned}$$

El primer terme de la norma al quadrat no depèn de b_i , per tant podrà sortir de l'integral factoritzant l'exponencial. Ara integrem el segon terme de la norma al quadrat, que no és una constant, ja que, és la funció de la densitat normal multivariant:

$$\begin{aligned}
 & \frac{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}} \int \frac{\exp\left[-(b_i - \hat{b}_i)^T \tilde{Z}_i^T \tilde{Z}_i (b_i - \hat{b}_i) / 2s^2\right]}{(2ps^2)^{q/2}} db_i = \\
 & = \frac{1}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}} \int \frac{\exp\left[-(b_i - \hat{b}_i)^T \tilde{Z}_i^T \tilde{Z}_i (b_i - \hat{b}_i) / 2s^2\right]}{(2ps^2)^{q/2} / \sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}} db_i = \\
 & = \frac{1}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}} = \frac{1}{\sqrt{|\tilde{Z}_i^T Z_i + \Delta^T \Delta|}}
 \end{aligned}$$

Per tant, la integral la podem expressar com:

$$\begin{aligned}
 & \int \frac{\exp\left(-\|\tilde{y}_i - \tilde{X}_i \mathbf{b} - \tilde{Z}_i b_i\|^2 / 2s^2\right)}{(2ps^2)^{q/2}} db_i = \\
 & = \int \frac{\exp\left(-\|\tilde{y}_i - \tilde{X}_i \mathbf{b} - \tilde{Z}_i b_i\|^2 / 2s^2\right)}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}} db_i
 \end{aligned}$$

Aleshores, tenim:

$$L(\mathbf{b}, \mathbf{J}, s^2 / y) = \frac{1}{(2ps^2)^{N/2}} \exp\left(\frac{-\sum_{i=1}^M \|\tilde{y}_i - \tilde{X}_i \mathbf{b} - \tilde{Z}_i \hat{b}_i\|^2}{2s^2}\right) \prod_{i=1}^M \frac{abs|\Delta|}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}} \quad (3)$$

Aquesta expressió la podem utilitzar per calcular el Màxim Likelihood i així estimar \mathbf{b}, \mathbf{J} i s^2 .

L'optimització és més simple si primer prenem la funció Likelihood com si fos una funció de \mathbf{J} . Calculem els estimadors condicionals $\hat{\mathbf{b}}(\mathbf{J})$ i $\hat{s}^2(\mathbf{J})$ que fa que $L(\mathbf{b}, \mathbf{J}, s^2)$ sigui màxim donat \mathbf{J} .

En la funció tenim una part que involucra a \mathbf{b} i \mathbf{s} i que es pot calcular per regressió lineal. La solució per mínims quadrats és:

$$(\hat{b}_1^T, \dots, \hat{b}_M^T, \hat{\mathbf{b}}^T)^T = \arg_{b_1, \dots, b_m, \mathbf{b}} \min \|y_e - X_e(b_1, \dots, b_M, \mathbf{b})^T\|^2$$

on

$$X_e = \begin{bmatrix} Z_1 & 0 & \dots & \dots & \dots & 0 & X_1 \\ \Delta & 0 & \dots & \dots & \dots & 0 & 0 \\ 0 & Z_2 & 0 & \dots & \dots & 0 & X_2 \\ \vdots & \Delta & 0 & & \dots & \dots & 0 \\ \vdots & 0 & \dots & \dots & \dots & 0 & \vdots \\ \vdots & \vdots & \dots & \dots & \dots & Z_M & X_M \\ 0 & 0 & \dots & \dots & \dots & \Delta & 0 \end{bmatrix} \text{ i } y_e = \begin{bmatrix} y_1 \\ 0 \\ y_2 \\ 0 \\ \vdots \\ y_M \\ 0 \end{bmatrix}$$

Podem escriure $(\hat{b}_1^T, \dots, \hat{b}_M^T, \hat{\mathbf{b}}^T)^T = (X_e^T X_e)^{-1} X_e^T y_e$, si aïllem podem treure l'estimador per $\hat{\mathbf{b}}(\mathbf{J})$.

La teoria de la regressió és donar l'estimador \mathbf{s}^2 del màxim condicional Likelihood.

$$\mathbf{s}^2(\mathbf{J}) = \frac{\|y_e - X_e(\hat{b}_1^T, \dots, \hat{b}_M^T, \hat{\mathbf{b}}^T)^T\|^2}{N} \quad (\text{suma de quadrats residual})$$

Si substituïm $\hat{\mathbf{b}}(\mathbf{J})$ i $\hat{\mathbf{s}}^2(\mathbf{J})$ a la funció Likelihood $L(\mathbf{J})$ ens queda:

$$L(\mathbf{J}) = L(\hat{\mathbf{b}}(\mathbf{J}), \mathbf{J}, \hat{\mathbf{s}}^2(\mathbf{J})) = \frac{\exp(-N/2)}{[2p\hat{\mathbf{s}}^2(\mathbf{J})]^{N/2}} \prod_{i=1}^M \frac{abs|\Delta|}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}}$$

Computacionalment, per calcular aquest estadístic, el paquet nlme de R fa les següents iteracions:

- Donat \mathbf{J} , calculem $\hat{\mathbf{b}}$ i $\hat{\mathbf{b}}(\mathbf{J})$.
- Un cop calculats aquest paràmetres, obtenim $\hat{\mathbf{s}}^2(\mathbf{J})$.
- Optimitzem la funció $L(\hat{\mathbf{b}}(\mathbf{J}), \mathbf{J}, \hat{\mathbf{s}}^2(\mathbf{J}))$ amb els paràmetres que hem calculat anteriorment, i obtenim \mathbf{J}' .
- Mirem si l'estadístic convergeix, i sinó tornem al pas inicial.

A partir de la densitat marginal $p(y_i / \mathbf{b}, \mathbf{J}, \mathbf{s}^2)$ podem expressar la densitat i derivar el Likelihood. Es pot escriure com una distribució normal de mitjana 0 i matriu de variàncies i covariàncies Σ .

El model el podem representar com:

$$y_i = X_i \mathbf{b} + Z_i \mathbf{b}_i + \mathbf{e}_i = X_i \mathbf{b} + \mathbf{e}_i^* \quad \text{on } i = 1, \dots, M$$

\mathbf{e}_i^* és la suma $(Z_i \mathbf{b}_i + \mathbf{e}_i)$ de dos normals multivariants independents de mitjana 0 i matriu de variàncies i covariàncies $\mathbf{s}^2 \Sigma_i$ on $\Sigma_i = Id + Z_i \Psi Z_i^T / \mathbf{s}^2$.

Podem dir que y_i és una normal multivariant de mitjana $X_i \mathbf{b}$ i variàncies i covariàncies $\mathbf{s}^2 \Sigma_i$:

$$p(y_i / \mathbf{b}, \mathbf{J}, \mathbf{s}^2) = (2\pi \mathbf{s}^2)^{-n/2} \exp\left(\frac{(y_i - X_i \mathbf{b})^T \Sigma_i^{-1} (y_i - X_i \mathbf{b})}{-2\mathbf{s}^2}\right) |\Sigma_i|^{-1/2}$$

Donat \mathbf{J} , els valors \mathbf{b} i \mathbf{s}^2 (màx. Likelihood) són:

$$\hat{\mathbf{b}}(\mathbf{J}) = \left(\sum_{i=1}^M X_i^T \Sigma_i^{-1} X_i \right)^{-1} \sum_{i=1}^M X_i^T \Sigma_i^{-1} y_i$$

$$\mathbf{s}^2(\mathbf{J}) = \frac{\sum_{i=1}^M (y_i - X_i \hat{\mathbf{b}}(\mathbf{J}))^T \Sigma_i^{-1} (y_i - X_i \hat{\mathbf{b}}(\mathbf{J}))}{N}$$

La **descomposició QR** d'una matriu triangular és el mètode numèric preferit per resoldre els problemes de mínims quadrats, s'escriu:

$$X = Q \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_t R$$

on X és una matriu $n \times p$ ($n \geq p$) de rang p , Q una matriu $n \times n$ ortogonal, R és $p \times p$ i triangular superior, la matriu Q_t és una matriu amb les p primeres columnes de Q . Diem que Q és ortogonal perquè $Q^T Q = Q Q^T = I$, això implica $Q_t^T Q_t = I$.

Una de les propietats més importants de les matrius ortogonals és que conserven la norma dels vectors després de multiplicar-los per Q o per Q^T .

Si apliquem Q^T al vector de residus dels mínims quadrats obtenim:

$$\begin{aligned}\|y - X\mathbf{b}\|^2 &= \|Q^T(y - X\mathbf{b})\|^2 = \|Q^T y - Q^T X\mathbf{b}\|^2 = \left\|c - Q^T Q \begin{bmatrix} R \\ 0 \end{bmatrix} \mathbf{b}\right\|^2 = \\ &= \left\|c - \begin{bmatrix} R \\ 0 \end{bmatrix} \mathbf{b}\right\|^2 = \|c_1 - R\mathbf{b}\|^2 + \|c_2\|^2\end{aligned}$$

on $c = (c_1^T c_2^T)^T = Q^T y$.

La solució per mínims quadrats de $\hat{\mathbf{b}}$ és fàcil avaluant $R\hat{\mathbf{b}} = c_1$ i la suma de quadrats residuals és $\|c_2\|^2$.

Aquesta descomposició la podem aplicar a la matriu augmentada del model \tilde{Z}_i :

$$\tilde{Z}_i = Q_{(i)} \begin{bmatrix} R_{11(i)} \\ 0 \end{bmatrix} \text{ on } Q_{(i)} \text{ és una matriu } (n_i + q) \times (n_i + q) \text{ i } R_{11(i)} \text{ és } (q \times q).$$

Aleshores,

$$\begin{aligned}\|\tilde{y}_i + \tilde{X}_i \mathbf{b} - \tilde{Z}_i \mathbf{b}_i\|^2 &= \|Q_{(i)}^T (\tilde{y}_i + \tilde{X}_i \mathbf{b} - \tilde{Z}_i \mathbf{b}_i)\|^2 = \\ &= \|c_{1(i)} - R_{10(i)} \mathbf{b} - R_{11(i)} \mathbf{b}_i\|^2 + \|c_{0(i)} - R_{00(i)} \mathbf{b}\|^2\end{aligned}$$

on $R_{10(i)}$ ($q \times p$), $R_{00(i)}$ ($n_i \times p$), el q -vector $c_{1(i)}$ i el n_i -vector $c_{0(i)}$ es defineixen per:

$$\begin{bmatrix} R_{10(i)} \\ R_{00(i)} \end{bmatrix} = Q_{(i)}^T \tilde{X}_i \text{ i } \begin{bmatrix} c_{1(i)} \\ c_{0(i)} \end{bmatrix} = Q_{(i)}^T \tilde{y}_i$$

Una altra manera de pensar és com a components de la descomposició ortogonal triangular de la matriu augmentada:

$$\begin{bmatrix} Z_i & X_i & y_i \\ \Delta & 0 & 0 \end{bmatrix} = Q_{(i)} \begin{bmatrix} R_{11(i)} & R_{10(i)} & c_{1(i)} \\ 0 & R_{00(i)} & c_{0(i)} \end{bmatrix}$$

Si tornem a la integral **(2)**, i substituïm, tenim que:

$$\begin{aligned}
 &= \int \frac{\exp\left(-\|y_i - X_i \mathbf{b} - Z_i b_i\|^2 + \|\Delta b_i\|^2 / 2\mathbf{s}^2\right)}{(2p\mathbf{s}^2)^{q/2}} db_i = \\
 &= \exp\left[\frac{\|c_{0(i)} - R_{00(i)} \mathbf{b}\|^2}{-2\mathbf{s}^2}\right] \int \frac{\exp\left(\frac{\|c_{1(i)} - R_{10(i)} \mathbf{b} - R_{11(i)} b_i\|^2}{-2\mathbf{s}^2}\right)}{(2p\mathbf{s}^2)^{q/2}} db_i
 \end{aligned}$$

La matriu $R_{11(i)}$ no és singular, per tant, podem aplicar el següent canvi de variable:

$$\mathbf{f}_i = (c_{1(i)} - R_{10(i)} \mathbf{b} - R_{11(i)} b_i / \mathbf{s}) \text{ i el diferencial } d\mathbf{f}_i = \mathbf{s}^{-q} \text{abs}[R_{11(i)}] db_i$$

Així, la integral ens queda de la següent manera:

$$\begin{aligned}
 &\int \frac{\exp\left(-\|c_{1(i)} - R_{10(i)} \mathbf{b} - R_{11(i)} b_i\|^2 / 2\mathbf{s}^2\right)}{(2p\mathbf{s}^2)^{q/2}} db_i = \\
 &= \frac{1}{\text{abs}[R_{11(i)}]} \int \frac{\exp\left(-\|\mathbf{f}_i\|^2 / 2\right)}{(2p)^{q/2}} d\mathbf{f}_i = \frac{1}{\text{abs}[R_{11(i)}]}
 \end{aligned}$$

Podem veure que és el mateix resultat que a **(3)** perquè:

$$\begin{aligned}
 \sqrt{|\tilde{Z}_i^T \tilde{Z}_i|} &= \sqrt{\left[R_{11(i)}^T \quad 0 \right] Q_{(i)}^T Q_{(i)} \begin{bmatrix} R_{11(i)} \\ 0 \end{bmatrix}} = \\
 &= \sqrt{|R_{11(i)}^T R_{11(i)}|} = \sqrt{|R_{11(i)}^T| |R_{11(i)}|} = \\
 &= \sqrt{|R_{11(i)}^T|^2} = \text{abs}[R_{11(i)}]
 \end{aligned}$$

Prenent aquest resultat i substituint, la funció Likelihood la podem escriure com:

$$\begin{aligned}
 L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2 / y) &= \prod_{i=1}^M \frac{\exp\left[-\|c_{0(i)} - R_{00(i)} \mathbf{b}\|^2 / 2\mathbf{s}^2\right]}{(2p\mathbf{s}^2)^{n_i/2}} \text{abs}\left(\frac{|\Delta|}{R_{11(i)}}\right) = \\
 &= \frac{\exp\left[-\sum_{i=1}^M \|c_{0(i)} - R_{00(i)} \mathbf{b}\|^2 / 2\mathbf{s}^2\right]}{(2p\mathbf{s}^2)^{N/2}} \prod_{i=1}^M \text{abs}\left(\frac{|\Delta|}{R_{11(i)}}\right)
 \end{aligned}$$

on el terme de l'exponencial té forma de suma de quadrats residual per \mathbf{b} .

Podem fer una segona descomposició ortogonal triangular:

$$\begin{bmatrix} R_{00(1)} & c_{0(1)} \\ \vdots & \vdots \\ R_{00(M)} & c_{0(M)} \end{bmatrix} = Q_0 \begin{bmatrix} R_{00} & c_0 \\ 0 & c_{-1} \end{bmatrix}$$

La funció es redueix a:

$$L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2 / y) = (2p\mathbf{s}^2)^{-N/2} \exp\left(\frac{\|c_{-1}\|^2 + \|c_0 - R_{00}\mathbf{b}\|^2}{-2\mathbf{s}^2}\right) \prod_{i=1}^M \text{abs}\left(\frac{|\Delta|}{|R_{11(i)}|}\right).$$

Donat \mathbf{q} , els valor per \mathbf{b} i \mathbf{s}^2 que maximitzen la funció de Likelihood són

$$\begin{aligned} \hat{\mathbf{b}}(\mathbf{q}) &= R_{00}^{-1} c_0 \\ \hat{\mathbf{s}}^2(\mathbf{q}) &= \frac{\|c_{-1}\|^2}{N} \end{aligned}$$

Tenint en compte aquests valors, Likelihood es defineix:

$$L(\mathbf{q} / y) = L(\hat{\mathbf{b}}(\mathbf{q}), \mathbf{q}, \hat{\mathbf{s}}^2(\mathbf{q}) / y) = \left(\frac{N}{2p\|c_{-1}\|^2}\right)^{N/2} \exp\left(-\frac{N}{2}\right) \prod_{i=1}^M \text{abs}\left(\frac{|\Delta|}{|R_{11(i)}|}\right)$$

Si prenem logaritmes obtenim:

$$\ell(\mathbf{q} / y) = \log L(\mathbf{q} / y) = \frac{N}{2} (\log N - \log(2p) - 1) - N \log \|c_{-1}\| + \sum_{i=1}^M \log \text{abs}\left(\frac{|\Delta|}{|R_{11(i)}|}\right)$$

Els paràmetres dels efectes aleatoris es poden avaluar utilitzant la matriu de descomposició ortogonal triangular:

$$\hat{b}_i(\mathbf{q}) = R_{11(i)}^{-1} (c_{1(i)} - R_{10(i)} \hat{\mathbf{b}}(\mathbf{q}))$$

➤ Restricted Likelihood

Quan estimem els components de la variància (\mathbf{s}^2 i \mathbf{s}_b^2) amb el **màxim Likelihood** s'acostuma a sobreestimar, molts analistes prefereixen utilitzar el **Restricted maximum likelihood (REML)**.

Existeixen moltes vies per definir el REML. Una manera és la de Laird and Ware, 1982:

$$L_R(\mathbf{J}, \mathbf{s}^2 / y) = \int L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2 / y) d\mathbf{b}$$

Integrem dins d'un context Bayesià, això significa que assumim a priori una distribució local uniforme pels efectes fixes (\mathbf{b}) i els integrem fora de la versemblança.

Utilitzant la norma del quadrat i el canvi de variable utilitzat anteriorment obtenim el log-restricted likelihood:

$$\begin{aligned} l_R(\mathbf{J}, \mathbf{s}^2 / y) &= \log L_R(\mathbf{J}, \mathbf{s}^2 / y) = \\ &= -\frac{N-P}{2} \log(2p\mathbf{s}^2) \frac{\|C_{-1}\|^2}{2\mathbf{s}^2} - \log \text{abs}|R_{00}| + \sum_{i=1}^M \log \text{abs} \left(\frac{|\Delta|}{|R_{11(i)}|} \right) \end{aligned}$$

Aconsegim l'estimador condicional $\mathbf{s}_R^2(\mathbf{J}) = \|C_{-1}\| / (N-P)$ per \mathbf{s}^2 , d'aquí obtenim:

$$\begin{aligned} l_R(\mathbf{J} / y) &= l_R(\mathbf{J}, \hat{\mathbf{s}}_R^2(\mathbf{J}) / y) = \\ &= \text{const} - (N-P) \log \|C_{-1}\| - \log \text{abs}|R_{00}| + \sum_{i=1}^M \log \text{abs} \left(\frac{|\Delta|}{|R_{11(i)}|} \right) \end{aligned}$$

Els components del log-restricted likelihood són molt similars al log-likelihood, excepte en què el log de la norma dels residus té un multiplicador diferent, i hi ha un component extra en el determinant.

$$\log \text{abs}|R_{00}| = \log \left| \sum_{i=1}^M X_i^T \Sigma_i^{-1} X_i \right| / 2$$

3.5. Referència a la llibreria utilitzada

Per estimar les variables i fer inferència, utilitzem la llibreria **nlme** del programa estadístic **R**. Existeixen altres programes per l'anàlisi dels models lineals mixtos per a mesures repetides com és el programa estadístic **SAS** amb el procediment **MIXED**. Hem optat per utilitzar el programa R perquè la metodologia de treball és molt semblant a la del programa SAS, però a diferència d'aquest, R és un programa gratuït.

La llibreria **nlme** és específica pels models lineals i no lineals mixtos, la principal funció que utilitzem és la **lme**. La llibreria **nlme** ha estat creada per José C. Pinheiro i Douglas M. Bates. Podem trobar més informació sobre l'ús d'aquesta llibreria al llibre *Mixed-effect Models in S and S-Plus* (2000)¹.

A continuació explicarem de manera breu els mètodes utilitzats i la forma de comparar els models ajustats.

I. Ajustar els models lineals mixtos

Tenim diferents funcions per ajustar els models lineals:

- **lm**: ajusta un model lineal simple.
- **lmList**: ajusta un model lm per a cada observació.
- **lme**: ajusta un model lineal mixt.

A continuació detallem els mètodes que es poden usar en cadascuna de les funcions.

i. Ajustar models lineals amb lm i lmList

El programa R té varies funcions i mètodes per ajustar els models lineals. La funció **lm** ajusta un model lineal general. Els mètodes que permet aquesta funció són:

- **print**: ens retorna la mínima informació dels resultats ajustat.
- **summary**: obtenim els resultats ajustats de manera més detallada.
- **plots**: per veure la qualitat de l'ajust.
- **predict**: obtenim les prediccions.
- **fitted/residuals/coef**: per extraure components.
- **anova**: comparem dos models.

La funció **lmList** ens ajusta un model **lm** per a cada grup o observació. Els mètodes més importants que permet aquesta funció són:

- **augPred**: augmentem les prediccions amb els valors observats.
- **coef**: ens retorna els coeficients per a cadascun dels models.
- **fitted**: ens dona els valors ajustats per a cada model.
- **fixef**: mitjana dels coeficients de cada model lm.

- **intervals**: intervals de confiança dels coeficients.
- **lme**: models lineals mixtos del **lmList** ajustat.
- **logLik**: suma dels log-likelihood de cada model (**lm**).
- **pairs**: diagrama de dispersió de la matriu de coeficients o dels efectes aleatoris.
- **plot**: s'utilitza per fer diagnòstics.
- **predict**: prediccions per a cada model ajustat.
- **print**: mínima informació sobre l'ajust.
- **qqnorm**: gràfic de normalitat.
- **ranef**: desviació de la mitjana dels coeficients.
- **resid**: residus de l'ajust per a cada model **lm** ajustat.
- **summary**: informació detallada sobre l'ajust **lm**.
- **update**: guardem els valors ajustats dels individus.

El principal propòsit de l'anàlisi preliminar amb **lmList** és donar una indicació de l'estructura dels efectes aleatoris per utilitzar en els models lineals mixtos. El mètode **pairs** ens proporciona una visió de l'estructura dels efectes aleatoris.

ii. Estimació dels models lineals mixtos amb lme

Utilitzem la funció **lme** per estimar els models lineals simples amb una estructura general de covariàncies pels efectes aleatoris.

a. Ajust del model lineal mixt simple

Utilitzem la funció **lme** per ajustar els models lineals mixtos amb màxim likelihood o el màxim likelihood restrictiu. Els mètodes més importants que permet, són similars a **lmList**:

- **ACF**: funció d'autocorrelació empírica dels residus.
- **anova**: test ràtio likelihood o test condicional.
- **augPred**: augmentem les prediccions amb els valors observats.
- **coef**: coeficients estimats pels diferents nivells.

- **fitted**: valors ajustats.
- **fixef**: estimacions dels efectes fixes.
- **intervals**: intervals de confiança dels paràmetres dels models.
- **logLik**: convergència del log-likelihood.
- **pairs**: diagrama de dispersió de la matriu de coeficients o dels efectes aleatoris.
- **plot**: s'utilitza per fer diagnòstics.
- **predict**: prediccions pels diferents grups de nivell.
- **print**: mínima informació sobre l'ajust.
- **qqnorm**: gràfic de normalitat.
- **ranef**: estimació dels efectes aleatoris.
- **resid**: residus pels diferents grups.
- **summary**: informació detallada sobre l'ajust .
- **update**: guardem els valors ajustats dels individus.

b. Discrepàncies en la matriu de variància – covariància

En un principi no considerem cap estructura en la matriu de la variància i covariància, però assumim que els efectes aleatoris són independents, per tant, la matriu \mathbf{y} és diagonal o bé, tenen la mateixa variància i la matriu està multiplicada per la matriu identitat.

A la llibreria nlme hi ha moltes classes de matrius definides positives anomenades **pdMat** classes. Aquesta funció s'utilitza en la matriu de variàncies i covariàncies dels efectes aleatoris. La funció per crear un objecte s'anomena constructor.

Els pdMat constructors tenen el mateix nom que la classe:

pdSymm: matriu simètrica definida positiva (és l'opció per defecte).

$$\Psi = \mathbf{s}^2 D = \begin{bmatrix} \mathbf{s}^2 d_{11} & \mathbf{s}^2 d_{12} & \cdots & \cdots & \mathbf{s}^2 d_{1q} \\ \mathbf{s}^2 d_{12} & \mathbf{s}^2 d_{22} & \cdots & \cdots & \mathbf{s}^2 d_{2q} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{s}^2 d_{1q} & \mathbf{s}^2 d_{2q} & \cdots & \cdots & \mathbf{s}^2 d_{qq} \end{bmatrix}$$

pdIdent: matriu diagonal definida positiva amb efectes aleatoris independents però amb la mateixa variància.

$$\Psi = \mathbf{s}^2 D = \begin{bmatrix} \mathbf{s}^2 & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{s}^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \mathbf{s}^2 \end{bmatrix}$$

pdDiag: matriu diagonal definida positiva amb efectes aleatoris independents.

$$\Psi = \mathbf{s}^2 D = \begin{bmatrix} \mathbf{s}^2 d_{11} & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{s}^2 d_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \mathbf{s}^2 d_{qq} \end{bmatrix}$$

pdCompSymm: matriu simètrica definida positiva amb efectes aleatoris independents però amb la mateixa variància i covariància.

$$\Psi = \mathbf{s}^2 D = \begin{bmatrix} \mathbf{s}^2 d_1 & \mathbf{s}^2 d_2 & \cdots & \cdots & \mathbf{s}^2 d_2 \\ \mathbf{s}^2 d_2 & \mathbf{s}^2 d_1 & \mathbf{s}^2 d_2 & \cdots & \mathbf{s}^2 d_2 \\ \vdots & \mathbf{s}^2 d_2 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{s}^2 d_2 \\ \mathbf{s}^2 d_2 & \mathbf{s}^2 d_2 & \cdots & \mathbf{s}^2 d_2 & \mathbf{s}^2 d_1 \end{bmatrix}$$

pdBlocked: matriu diagonal amb efectes aleatoris amb subgrups independents.

$$D = \begin{bmatrix} D_{11} & 0 & 0 & \cdots & 0 \\ 0 & D_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & D_{qq} \end{bmatrix}$$

Els valors inicials de \mathbf{y} poden ser derivats de la funció lme, els **pdMat** es poden utilitzar únicament per especificar una classe i una formula pels efectes aleatoris.

c. Ajust dels models multinivell

L'ajust dels models multinivell és igual que pel model simple on cada nivell admet una estructura diferent.

iii. Examinem el model ajustat

Abans de començar a fer inferència hem de comprovar que la distribució normal que assumim sigui vàlida per a les nostres dades.

Tenim dos blocs de comprovacions:

a. Comprovació dels residus del model

Els errors són independents i idènticament distribuïts. Segueixen una normal de mitjana 0 i variància σ^2 ; i són independents dels efectes aleatoris. Per validar-ho utilitzem una sèrie de gràfics:

- **Resposta observada vers valors ajustats.** Si fos necessari, es podria afegir alguna covariable d'interès. A partir d'aquest gràfic podem veure si la hipòtesis de linealitat es compleix.
- **Valors ajustats vers residus estandarditzats:** comprovem que la variància sigui constant. Si el residu d'un dels valors observats és gran, ens indicarà que aquesta observació no és normal, pot ser un cas on la mesura estigui presa incorrectament o un cas anòmal.

En les nostres dades, a vegades s'observen patrons, això ens pot portar a pensar que aquest gràfic no és correcte, però el que passa és que les dades estan mesurades amb una precisió en concret.

- **Gràfic de normalitat (qqnorm)**: comprovem la normalitat dels residus.

b. Comprovació de les hipòtesis dels efectes aleatoris

Els efectes aleatoris segueixen una normal de mitjana 0 i variància σ^2 i són independents per grups diferents. Els mètodes utilitzats són:

- **gràfic de normalitat (qqnorm)**: comprovem si les dades segueixen una distribució normal, podem identificar outliers si les cues de la distribució són pesants.

- **gràfic de dispersió (pairs)**, comprovem la homogeneïtat dels efectes aleatoris (residus vers efectes aleatoris). A partir d'aquest gràfic observem si la variància dels efectes aleatoris és constant.

II. Extensió dels models lineals mixtos

Quan definim els models lineals mixtos suposem que els efectes aleatoris són independents i idènticament distribuïts de mitjana 0 i variància constant. Ens podem trobar que els errors dels efectes aleatoris siguin heteroscedàstics o que estiguin correlacionats o ambdues coses simultàniament.

La matriu de variàncies i covariàncies dels efectes aleatoris es pot descompondre en dos components: l'estructura de la variància i l'estructura de les correlacions.

i. Formulació general del model ampliat

En el model lineal amb efectes mixtos assumim que els errors dins del grup (\mathbf{e}_i) dels efectes aleatoris són independents $N(0, \sigma^2 \mathbf{I})$.

En el model ampliat assumim la heteroscedasticitat i la correlació dels errors dels efectes aleatoris:

$$y_i = X_i \mathbf{b} + Z_i b_i + \mathbf{e}_i$$

on $i = 1, \dots, M$, $b_i \sim N(0, \mathbf{y})$ i $\mathbf{e}_i \sim N(0, \mathbf{s}^2 \Lambda_i)$

on Λ_i és una matriu definida positiva parametritzada per un nombre fix, generalment petit, de paràmetre \mathbf{I} . En el model bàsic assumim que els errors entre grups \mathbf{e}_i són independents per diferents i ; i els efectes aleatoris b_i són independents. La \mathbf{s}^2 es factoritza a fora de la matriu Λ_i .

En el cas de tenir dos nivells, $\mathbf{e}_{ij} \sim N(0, \mathbf{s}^2 \Lambda_{ij})$ amb $i = 1, \dots, M$; $j = 1, \dots, M_i$ on Λ_{ij} és una matriu positiva i parametritzada per un nombre fix de vectors \mathbf{I} .

ii. Estimació i mètodes computacionals

El fet que Λ_i sigui positiva ens permet:

$$\Lambda_i = \left(\Lambda_i^{-1/2} \right)^T \left(\Lambda_i^{1/2} \right) \text{ o } \Lambda_i^{-1} = \left(\Lambda_i^{-1/2} \right)^T \left(\Lambda_i^{-1/2} \right)$$

Podem escriure:

$$y_i^* = \left(\Lambda_i^{-1/2} \right)^T y_i; \mathbf{e}_i^* = \left(\Lambda_i^{-1/2} \right)^T \mathbf{e}_i; X_i^* = \left(\Lambda_i^{-1/2} \right)^T X_i \text{ i } Z_i^* = \left(\Lambda_i^{-1/2} \right)^T Z_i$$

$$\text{aleshores } \mathbf{e}_i^* \sim N \left[\left(\Lambda_i^{-1/2} \right)^T \cdot 0, \mathbf{s}^2 \left(\Lambda_i^{-1/2} \right)^T \Lambda_i \Lambda_i^{-1/2} \right] = N(0, \mathbf{s}^2 \mathbf{I})$$

Podem reescriure el model com:

$$y_i^* = X_i^* \mathbf{b} + Z_i^* b_i + \mathbf{e}_i^* \quad i = 1, \dots, M \quad \text{on } b_i \sim N(0, \mathbf{y}), \mathbf{e}_i^* \sim N(0, \mathbf{s}^2 \mathbf{I}).$$

El diferencial de la transformació lineal $\left(y_i^* = \left(\Lambda_i^{-1/2} \right)^T y_i \right)$ és simple

$$dy_i^* = \left| \Lambda_i^{-1/2} \right| dy_i.$$

La funció likelihood del model ampliat s'expressa:

$$\begin{aligned} L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2, \mathbf{I} / y) &= \prod_{i=1}^M p(y_i / \mathbf{b}, \mathbf{J}, \mathbf{s}^2, \mathbf{I}) = \\ &= \prod_{i=1}^M p(y_i^* / \mathbf{b}, \mathbf{J}, \mathbf{s}^2, \mathbf{I}) \cdot \left| \Lambda_i^{-1/2} \right| = L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2, \mathbf{I} / y_i^*) \cdot \prod_{i=1}^M \left| \Lambda_i^{-1/2} \right| \end{aligned}$$

La funció $L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2, \mathbf{I} / y_i^*)$ correspon al model lineal bàsic.

Així, la funció Restricted likelihood pel model general, s'escriu:

$$L_R(\mathbf{J}, \mathbf{s}^2, \mathbf{I} / y) = \int L(\mathbf{b}, \mathbf{J}, \mathbf{s}^2, \mathbf{I} / y) d\mathbf{b} = L_R(\mathbf{J}, \mathbf{s}^2, \mathbf{I} / y^*) \prod_{i=1}^M \left| \Lambda_i^{-1/2} \right|$$

iii. Extensió del model lineal sense efectes aleatoris

La matriu de variància i covariància del vector resposta y_i en l'extensió del model lineal mixt és $\text{var}(y_i) = \Sigma_i = \mathbf{s}^2 (\mathbf{Z}_i \mathbf{D} \mathbf{Z}_i^T + \Lambda_i)$ té dos components. La component que s'utilitza per modelar la heteroscedasticitat i la correlació (efectes aleatoris) $\mathbf{Z}_i \mathbf{D} \mathbf{Z}_i^T$ i la component dins del grup Λ_i .

Aquestes dues components es poden modelar. En alguns aplicatius es pot evitar incorporar efectes arbitraris per explicar la dependència entre les observacions i utilitzar la component Λ_i per modelar l'estructura de la matriu de variàncies i covariàncies, d'aquesta manera obtenim una versió més simplificada dels models d'efectes mixtos:

$$y_i = X_i \mathbf{b} + \mathbf{e}_i \text{ on } \mathbf{e}_i \sim N(0, \mathbf{s}^2 \Lambda_i) \quad i = 1, \dots, M$$

Si utilitzem les mateixes transformacions que en el model extens obtenim la regressió lineal clàssica:

$$y_i^* = X_i^* \mathbf{b} + \mathbf{e}_i^* \text{ on } \mathbf{e}_i^* \sim N(0, \mathbf{s}^2 \mathbf{I}) \quad i = 1, \dots, M$$

Aquest model es pot abordar des de la rutina gls de la llibreria nlme del programa R.

iv. Descomposició de l'estructura de la matriu de variàncies i covariàncies dins del grup

La matriu Λ_i sempre es pot descompondre com un producte de dues matrius simples: $\Lambda_i = V_i C_i V_i$ on V_i és diagonal i C_i matrius de correlacions, matriu definida positiva i tots els elements diagonals iguals a 1. La matriu V_i no és única, en el sentit que nosaltres podem multiplicar qualsevol de les seves files per -1 i continuem tenint la mateixa descomposició. Per assegurar la unicitat de la descomposició, forcem que els elements de la diagonal de V_i siguin positius. És fàcil verificar que $Var(\mathbf{e}_{ij}) = \sigma^2 [V_i]_{jj}^2$ i $Cov(\mathbf{e}_{ij}, \mathbf{e}_{jk}) = [C_i]_{jk}$, per tant, V_i descriu la variància i C_i descriu la correlació intra-grup dels errors. Aquesta descomposició Λ_i amb l'estructura de la variància i correlacions és interessant tant en l'aspecte teòric com en el computacional.

v. Funcions de variància per modelar la heteroscedasticitat

Les funcions de variància s'utilitzen per modelar la matriu de variàncies dels errors intra-grup utilitzant covariables.

Definim el model general de la funció de variància pels errors intra-grups en el primer nivell lineal extens del model d'efectes mixtos com

$$Var(\mathbf{e}_{ij}|b_i) = \sigma^2 g^2(\mathbf{m}_{ij}, \mathbf{u}_{ij}, \mathbf{d}), \quad i = 1, \dots, M, \quad j = 1, \dots, n_i,$$

on $\mathbf{m}_{ij} = E[y_{ij}|b_i]$, \mathbf{u}_{ij} és un vector de variància de les covariables, \mathbf{d} és un vector de variància dels paràmetres i $g(\cdot)$ és la funció de variància, assumim que és continua a \mathbf{d} .

La funció de variància d'un sol nivell es pot generalitzar a models multinivell. Per exemple, la funció de variància per a un model de dos nivells és:

$$\text{Var}(\mathbf{e}_{ijk} | b_{i,j}, b_{ij}) = \mathbf{S}^2 g^2(\mathbf{m}_{ijk}, \mathbf{u}_{ijk}, \mathbf{d}),$$

on $i = 1, \dots, M$, $j = 1, \dots, M_i$, $k = 1, \dots, n_{ij}$, i $\mathbf{m}_{ijk} = E(y_{ijk} | b_{i,j}, b_{ij})$.

La funció de variància pel model lineal desenvolupat es defineix de forma semblant, però, com que no hi ha efectes aleatoris presents, i el model de la variància marginal no involucra cap aproximació, es pot expressar com:

$$\text{Var}(\mathbf{e}_{ij}) = \mathbf{S}^2 g^2(\mathbf{m}_{ij}, \mathbf{u}_{ij}, \mathbf{d}), \quad i = 1, \dots, M, \quad j = 1, \dots, n_i,$$

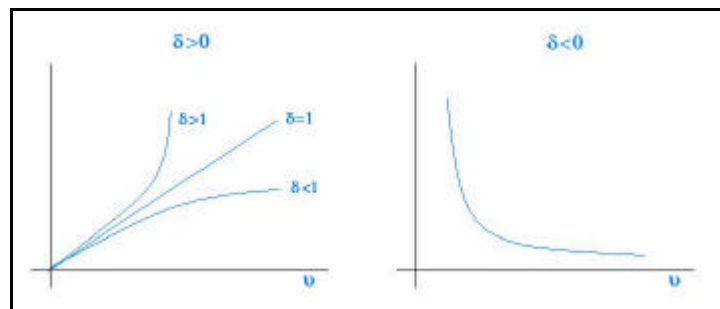
on $\mathbf{m}_{ij} = E[y_{ij}] = \mathbf{x}_{ij}^T \mathbf{b}$ i \mathbf{u}_{ij} és el vector de variància de les covariables i \mathbf{d} és el vector de variància dels paràmetres.

La llibreria **nlme** conté un col·lecció de classes de funcions de variància, la classe **varFunc**, que utilitzem per especificar la variància intragrupos del model en qualsevol dels models d'efectes mixtos lineals desenvolupats, o el model lineal desenvolupat.

Les principals funcions estàndards de la funció varFunc són:

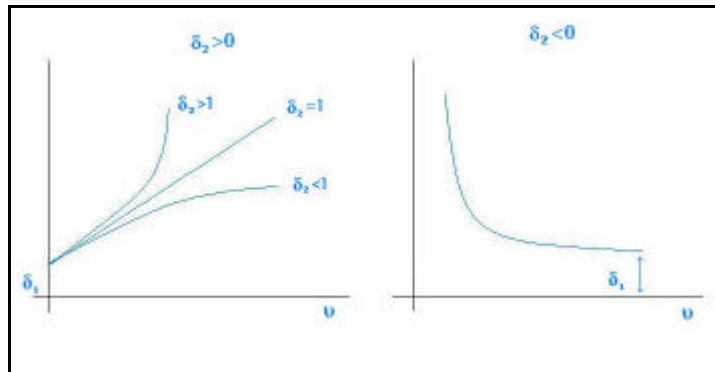
- **varFixed:** Variància fixa
- **varIdent:** Variància constant per grups
- **varPower:** Augmentem la variància amb una potència

$$(g_{pow}(\mathbf{u}, \mathbf{d}) = |\mathbf{u}|^d).$$



- **varExp:** Augmentem la variància de forma exponencial, és un cas concret de varPower. ($g_{exp}(\mathbf{u}, \mathbf{d}) = \exp(d\mathbf{u})$).

- **varConstPower:** Augmentem la variància amb una constant i una potència ($g_{cpow}(u, d) = d_1 + |u|^{d_2}$).



- **varComb:** Combinació de funcions de la variància

Les funcions de variància estan especificades dins la funció *lme* utilitzant l'argument *weights* (pesos). Per defecte, *weights* = *NULL*, correspon al model amb variància homoscedàstica per els errors intragrup.

vi. Estructures de correlació per modelar dependencies

Les estructures de correlació s'utilitzen per modelar dependencies entre observacions. En el context dels models mixtos i els model lineals extensos, s'utilitzen per modelar la dependència intragrup d'errors. Històricament la correlació d'estructures s'ha desenvolupat per dues classes principals de dades: series temporals i series a l'espai.

A la llibreria **nlme** tenim una sèrie de classes de correlació d'estructures, les classes **corStruct**, són utilitzades específicament per modelar la correlació intra-grups en qualsevol dels models lineals extensos d'efectes mixtos, o els models lineals extensos.

Les principals classes de la funció **corStruct** són:

- **corAR1:** Classe que implementa una correlació d'estructura autoregressiva d'ordre 1.

$$\mathbf{e}_t = r\mathbf{e}_{t-1} + u_t, \quad |r| < 1 \quad \text{on } u_t \text{ és soroll blanc}$$

$$C = \begin{bmatrix} 1 & r & r^2 & r^3 & \dots & r^n \\ r & 1 & r & r^2 & \dots & r^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r^n & r^{n-1} & \dots & r^2 & r & 1 \end{bmatrix}$$

- **corCAR1:** Classe que implementa una estructura continua temporal

$$\text{AR}(1). \text{cor}(\mathbf{e}_t, \mathbf{e}_s) = r^{|t-s|}, \quad 0 \leq r \leq 1$$

- **corARMA1:** Classe que s'utilitza per especificar una estructura autoregressiu de mitjanes mòbils pels errors intra-grups.

$$\mathbf{e}_t = u_t + \sum_{i=1}^p \Phi_i \mathbf{e}_{t-i} + \sum_{j=1}^q \theta_j \mathbf{e}_{t-j}$$

- **corSymm:** Classe que implementa una correlació general d'estructures.

$$C = \begin{bmatrix} 1 & r_{11} & r_{12} & \dots & r_{1n} \\ r_{11} & 1 & r_{21} & \dots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & r_{n-n} \\ r_{1n} & r_{2n} & \dots & r_{n-1n} & 1 \end{bmatrix}$$

- **corCompSymm:** Classe que implementa una correlació d'estructures de simetria composta.

$$C = \begin{bmatrix} 1 & r & \dots & \dots & r \\ r & 1 & r & \dots & r \\ \vdots & r & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & r \\ r & r & \dots & r & 1 \end{bmatrix}$$

- **corRatio:** Classe que implementa una estructura a l'espai representant una ratio quadràtica.

- **corExp:** Estructura a l'espai representant una distribució exponencial.

$$r(d) = \exp(-d/q)$$

- **corGaus:** Estructura a l'espai representant una distribució gaussiana.

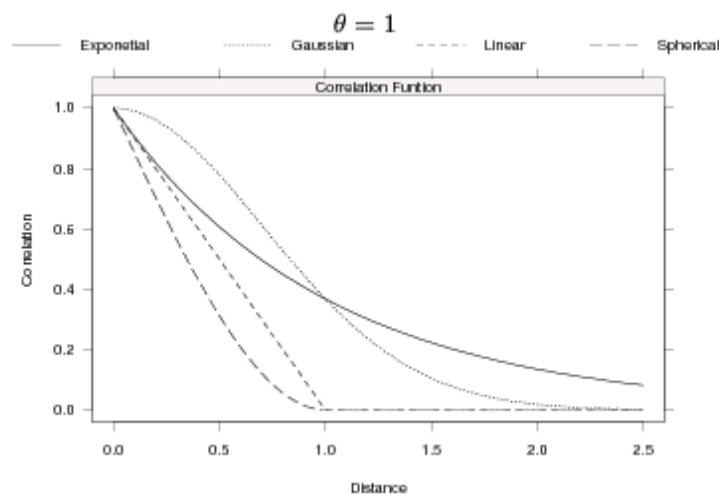
$$r(d) = \exp\left(-\left(d/q\right)^2\right)$$

- **corLin:** Estructura a l'espai representant una distribució lineal.

$$r(d) = (1 - d/q) \text{ on } (d \leq q)$$

- **corSpher:** Estructura a l'espai representant una distribució esfèrica.

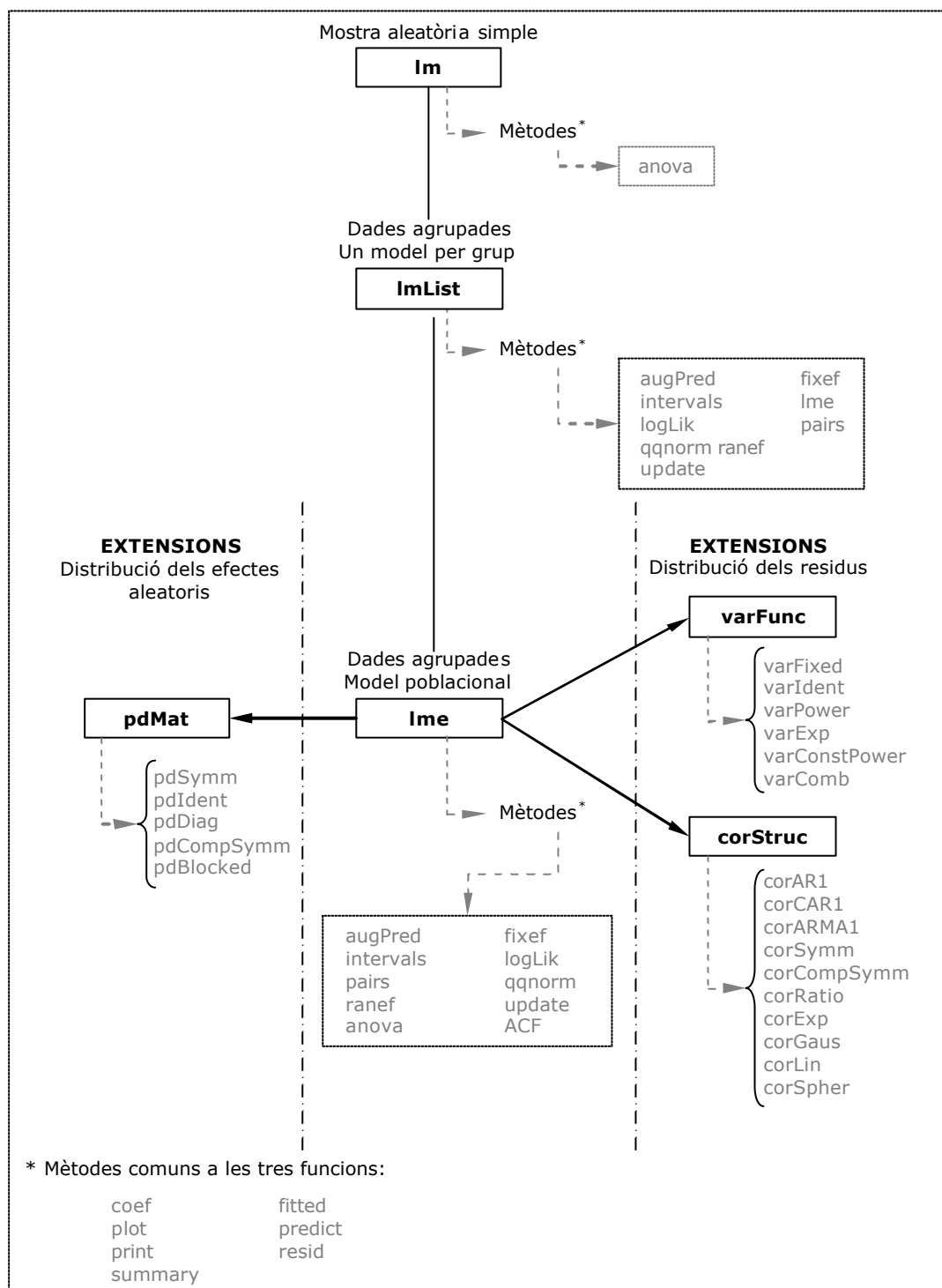
$$r(d) = \left(1 - 1.5(d/q) + 0.5(d/q)^3\right) \text{ on } (d \leq q)$$



Els dos arguments principals de la majoria dels constructors **corStruct** són *value* i *form*. El primer especifica els valors dels paràmetres de correlació i el segon és una formula parcial especificant el vector de posició i, opcionalment, una variable de grup per les dades (observacions en grups diferents s'assumeixen com a independents).

Les estructures de correlació estan especificades a la funció *lme* mitjançant l'argument de correlació (*correlation*). Per defecte, *correlation=NULL*, corresponent a errors intra-grups no correlacionats. Les estructures de correlació estan especificades com objectes **corStruct**, creats mitjançant els objectes estàndards descrits anteriorment, o un constructor **corStruct** definit per l'usuari.

III. Resum de la llibreria



4. Resultats

4.1. Base de dades

El fitxer original està format per 689 pacients (amb indicadors no correlatius) i 3064 registres (1 per a cada ull i visita). Les dades es recollien en una base de dades en format ACCESS. Aquesta base de dades està formada per dues taules, Pacients i Visites, que s'unien a partir de l'identificador del pacient. La metgessa visitava als pacients i era l'encarregada de mantenir actualitzades aquestes taules.

Base de dades inicial

La taula Pacients està formada per les següents variables:

- **Pacients:** És la variable que conté el número corresponent a l'identificador de cada pacient.
- **Sexe:** És la variable categòrica que ens indica el sexe del pacient. En aquest cas: v= Home i h=Dona.
- **Data de Naixement:** És la variable que ens indica la data de naixement del pacient.

La taula **Visites** conté les següents variables:

- **Pacients:** És la variable on trobem el número corresponent a l'identificador de cada pacient.
- **Data de la visita:** Aquesta variable ens indica el dia, mes i any en què es va realitzar la visita.
- **RefOD:** Variable numèrica que conté el valor de l'índex refractari de l'ull dret.
- **RefOI:** Variable numèrica que conté el valor de l'índex refractari de l'ull esquerre.
- **AstOD:** Variable numèrica que conté el valor d'astigmatisme de l'ull dret.
- **AstOI:** Variable numèrica que conté el valor d'astigmatisme de l'ull esquerre.

Per tal de facilitar el treball amb les dades, es van crear noves variables a partir de les esmentades anteriorment.

Base de dades seguiment

Aquesta base de dades conté les variables de la base de dades inicial i les noves variables calculades. Les variables calculades són:

- **Visita:** És la variable numèrica que conté l'ordinal de la visita. Aquesta variable ens indica longitudinalitat, però no l'utilitzem com a covariable principal perquè les dades no estan mesurades en els mateixos instants de temps en tots els pacients. Per exemple, un pacient en la seva segona visita pot ser més petit que un pacient en la seva primera visita.
- **N. Visites:** És la variable numèrica que conté el total de visites realitzades per un mateix pacient.
- **Mesos:** És la variable numèrica que ens indica l'edat del pacient, en mesos, en el moment de la visita. Aquesta variable es calcula a partir de la data de naixement i la data de la visita.
- **Anys:** És la variable numèrica que ens indica l'edat dels pacients, en anys, en el moment de la visita. Aquesta variable està calculada amb la data de naixement i la data de la visita.
- **MesIni:** És la variable numèrica que ens indica l'edat en mesos en el moment en què es va realitzar la primera visita.
- **MesFin:** És la variable numèrica que ens indica l'edat en mesos en el moment en què es va realitzar l'última visita.
- **Periodo:** És la variable numèrica que ens indica durant quants mesos es va visitar el pacient. Aquesta variable es calcula a partir de l'operació: $\text{MesFin} - \text{MesIni}$.
- **IndOD:** És la variable numèrica que conté el valor de l'índex refractari de l'ull dret.
- **IndOI:** És la variable numèrica que conté el valor de l'índex refractari de l'ull esquerre.
- **PrimIOD:** Aquesta variable numèrica conté el valor de l'equivalent esfèric inicial de l'ull dret.

- **PrimIOI:** Aquesta variable numèrica conté el valor de l'equivalent esfèric inicial de l'ull esquerre.
- **UltimIOD:** Variable numèrica que conté el valor de l'equivalent esfèric final de l'ull dret.
- **UltimIOI:** Variable numèrica que conté el valor de l'equivalent esfèric final de l'ull esquerre.
- **DifOD:** Variable numèrica que és el resultat de l'operació: $\text{UltimIOD} - \text{PrimIOD}$.
- **DifOI:** Variable numèrica que és el resultat de l'operació: $\text{UltimIOI} - \text{PrimIOI}$.
- **Tip1VOD:** Variable categòrica que ens explica la patologia que presenta l'ull dret en la primera visita del pacient: E=emmetropia, H=hipermetropia i M=miopia.
- **Tip1VOI:** Variable categòrica que ens explica la malaltia que presenta l'ull esquerre en la primera visita del pacient: E=emmetropia, H=hipermetropia i M=miopia.
- **ClasOD:** Variable categòrica que ens explica l'evolució de l'ull dret: D=decrement, S= estable i I =increment.
- **ClasOI:** Variable categòrica que ens explica l'evolució de l'ull esquerre: D=decrement, S=estable i I=increment.

Fitxer de treball pel programa R

Per poder treballar amb el programa R hem de crear un nou fitxer on cada línia representi una observació, és a dir, cada visita tindrà dos registres un per a cada ull.

- **id:** És la variable que conté el número corresponent a l'identificador de cada pacient.
- **Sexe:** És la variable categòrica que ens indica el sexe del pacient. En aquest cas: 0=Home i 1=Dona.
- **Visit:** És la variable numèrica que conté l'ordinal de la visita.
- **Nvisit:** És la variable numèrica que conté el total de visites realitzades per un pacient concret.
- **Mes:** És la variable numèrica que conté l'edat del pacient, en mesos, en el moment de la visita.

- **Ini:** És la variable numèrica que conté l'edat, en mesos, en el moment en què es va realitzar la primera visita.
- **Period:** És la variable numèrica que conté el període de mesos que un individu va estar a l'estudi.
- **Ee:** És la variable resposta, ens indica l'equivalent esfèric. L'equivalent esfèric és una suma algebraica del valor de l'esfera i la meitat del valor cilíndric, o el que és el mateix, l'indicador refractari més un mig de l'astigmatisme.
- **Eye:** És la variable categòrica que ens indica quin ull estem estudiant. En aquest cas: 0=dret i 1=esquerra.
- **Ast:** És la variable numèrica que conté el valor d'astigmatisme de l'ull.
- **Enf:** És la variable categòrica que ens indica la malaltia de l'ull estudiat. En aquest cas: 0=Emmetropia, -1=Miopia i 1=Hipermetropia.
- **Evol:** És la variable categòrica que ens indica l'evolució de la malaltia en l'ull estudiat. En aquest cas: 0=estable, -1=decreixent i 1=creixent.

4.2. Descriptius

Abans de començar a modelitzar les dades fem una breu descriptiva a les variables d'interès.

Per tal que fos més senzill realitzar els **descriptius de la primera visita** de cada individu, hem utilitzat el programa SPSS.

Dades Inicials

Tenim un total de 689 pacients (3064 registres o observacions) al fitxer de dades inicials.

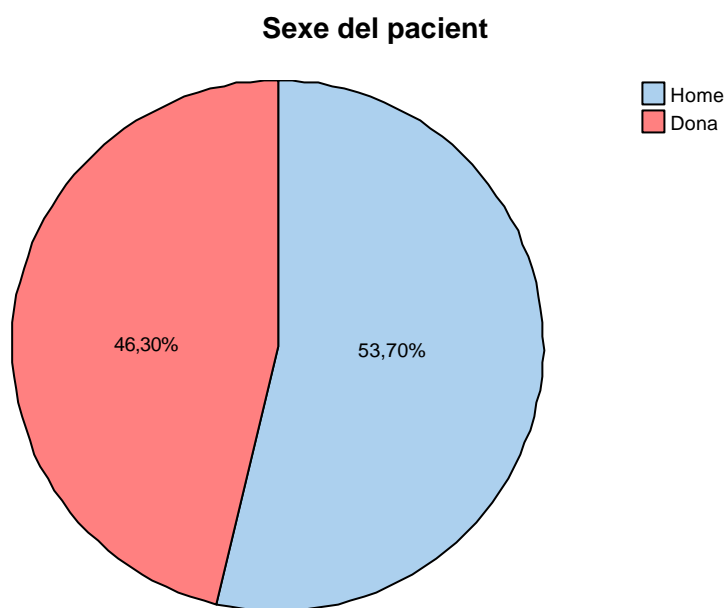


Figura 1: Gràfic del sexe dels pacients

En aquest primer fitxer, tenim un percentatge lleugerament superior d'homes respecte al de les dones (53.7% i 46.3% respectivament).

Sexe del pacient		Malaltia del pacient a la primera visita	
Home	370	Miopia	140
Dona	319	Emmetropia	340
		Hipermetropia	209

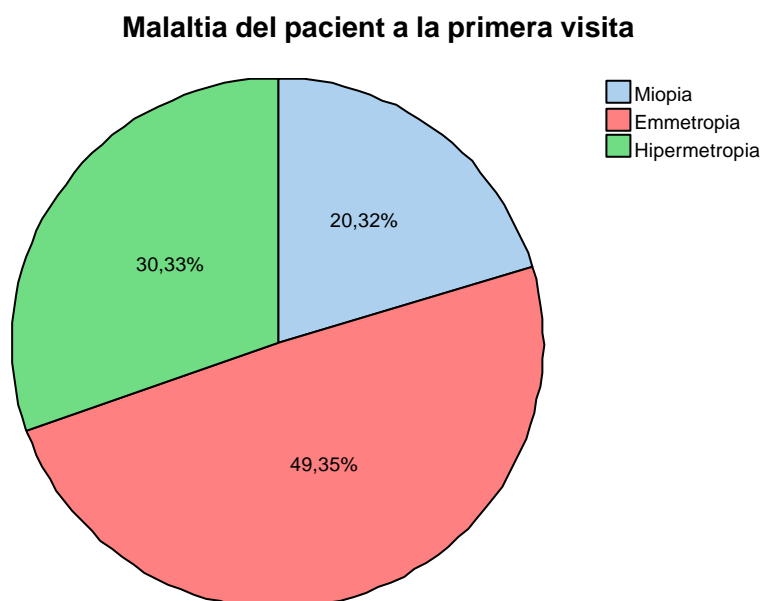


Figura 2: Malaltia del pacient en la primera visita

A la **figura 2**, observem que a la primera visita gairebé el 50% dels pacients són emmetrops. El 30% tenen hipermetropia i un 20% miopia.

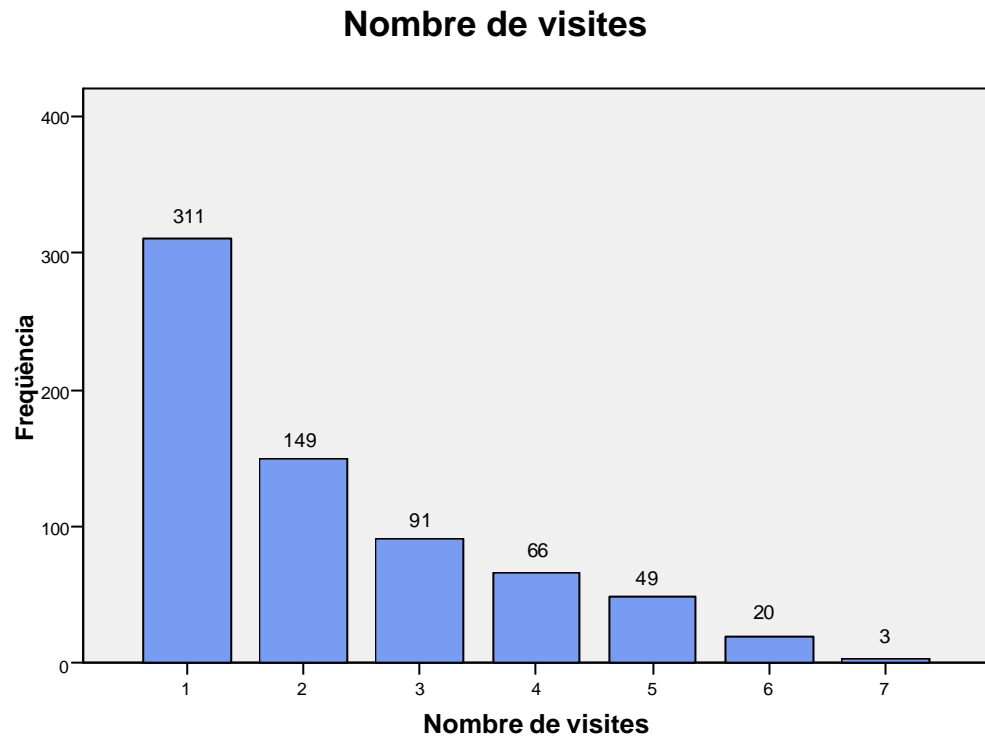


Figura 3: Gràfic del nombre de visites

A la taula anterior, veiem que tenim 311 pacients que només han tingut una visita. Aquests individus formen part del grup dels pacients que han tingut una evolució estable.

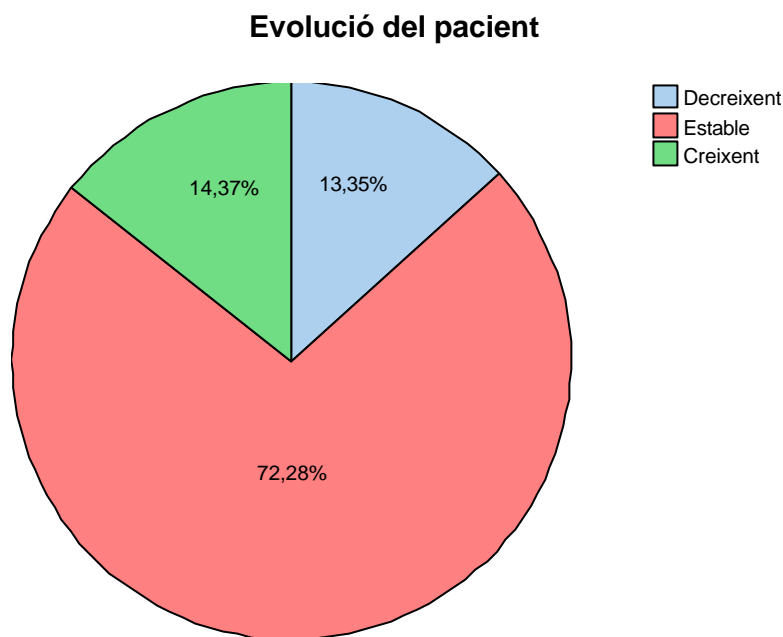


Figura 4: Evolució del pacient

Evolució del pacient	
Decreixent	92
Estable	498
Creixent	99

En la figura **4** podem veure que pràcticament tres quartes parts dels pacients es van mantenir estables durant el seguiment.

A la figura següent, podem veure els pacients amb astigmatisme a la primera visita.

En aquest gràfic observem que un 31.5% dels pacients (concretament 217) no tenen aquesta malaltia. També podem veure que la presició de la mesura de l'astigmatisme és de 0,25 en 0,25.

Astigmatisme

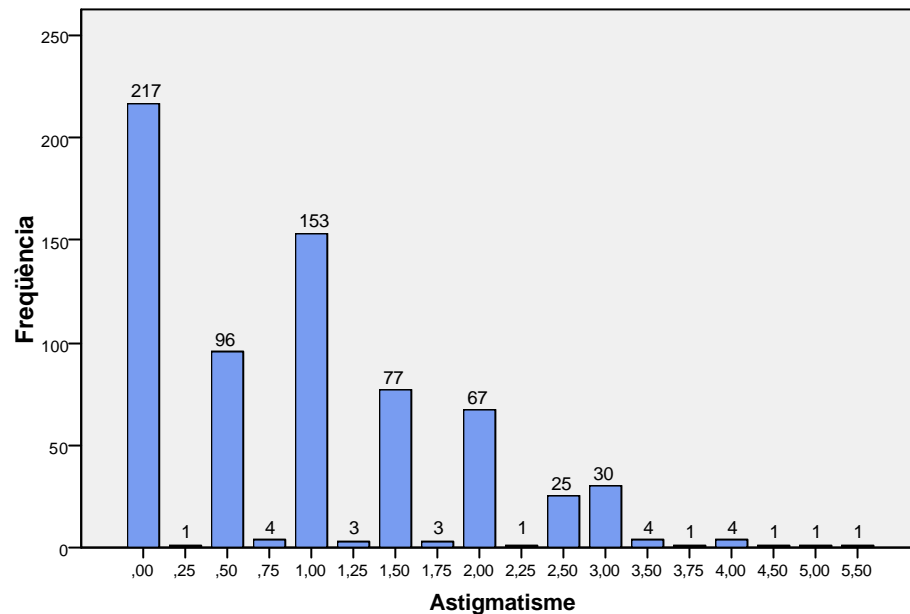


Figura 5: Gràfic d'astigmatisme

Edat a la primera visita		Equivalent esfèric	
Min.	1	Min.	-14
1st. Qu:	15,5	1st. Qu:	0
Median	42	Median	1,5
Mean	56,87	Mean	0,98168
3rd.Qu:	89	3rd.Qu:	3
Max.	280	Max.	11

Període de seguiment		Astigmatisme	
Min.	0	Min.	0
1st. Qu:	0	1st. Qu:	0
Median	20	Median	1
Mean	37,92	Mean	0,967
3rd.Qu:	70	3rd.Qu:	1,5
Max.	160	Max.	5,5

En el moment de la primera visita, aquests pacients tenien una edat compresa entre 1 mes i 280 mesos.

Podem veure que la distribució dels valors de l'equivalent esfèric es troba lleugerament desplaçada cap a la dreta, ja que tant la mitjana com la mediana tenen valors positius. Com hem dit anteriorment, tenim 311 pacients amb una sola visita. Aquests pacients tindran una evolució estable, sense cap tendència. Com que nosaltres volem estudiar la tendència d'aquestes malalties, eliminarem aquestes dades del nostre fitxer de treball final.

Per un motiu similar, eliminarem del fitxer els individus amb 2 visites, ja que, el model per les dades d'aquests pacients estimaria directament un model exacte. Si eliminem aquests individus de la mostra, obtindrem un model més robust.

Dades Finals

Un cop eliminats aquests individus, tenim un total de 229 pacients (1846 registres o observacions) amb una edat a la primera visita compresa entre 1 mes i 171 mesos. El període de seguiment va entre 18 mesos i 160 mesos.

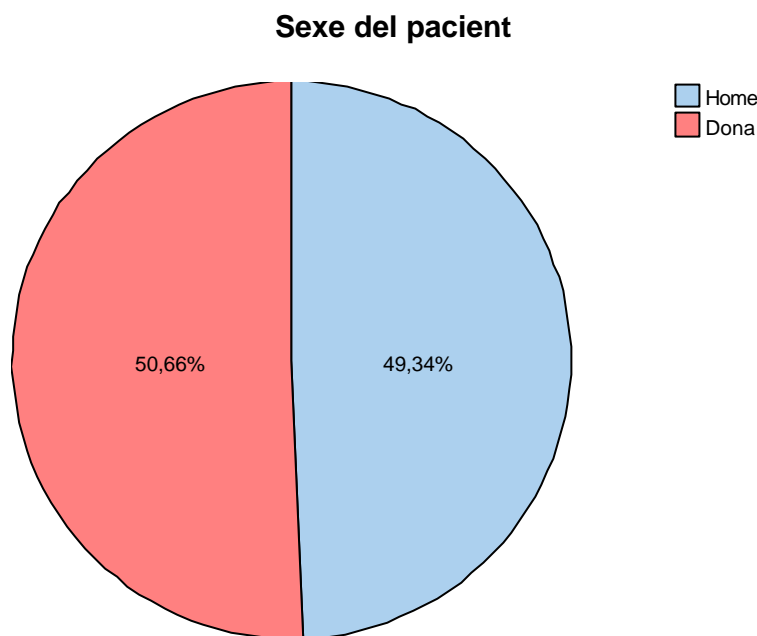
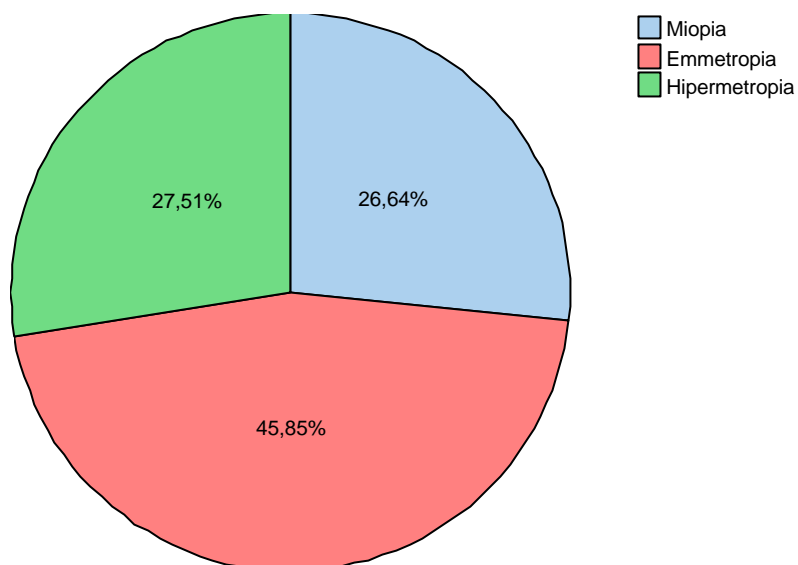


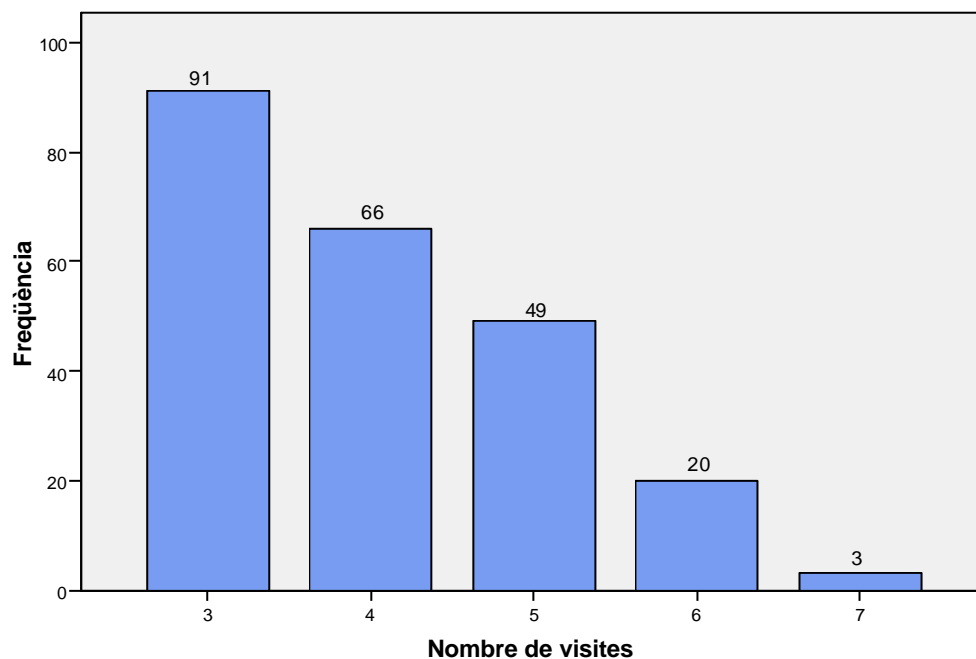
Figura 6: Gràfic del sexe dels pacients

Podem veure que a l'eliminar individus, la mida de la mostra s'ha igualat per sexes, un 49.3% d'homes i 50.7% de dones.

Sexe del pacient		Malaltia del pacient a la primera visita	
Home	113	Miopia	61
Dona	116	Emmetropia	105
		Hipermetropia	63

Malaltia del pacient a la primera visita**Figura 7: Malaltia del pacient en la primera visita**

Observem que els pacients amb poques visites eren emmetrops, ja que hem passat de tenir 340 pacients a 105 en aquesta categoria.

Nombre de visites**Figura 8: Gràfic del nombre de visites**

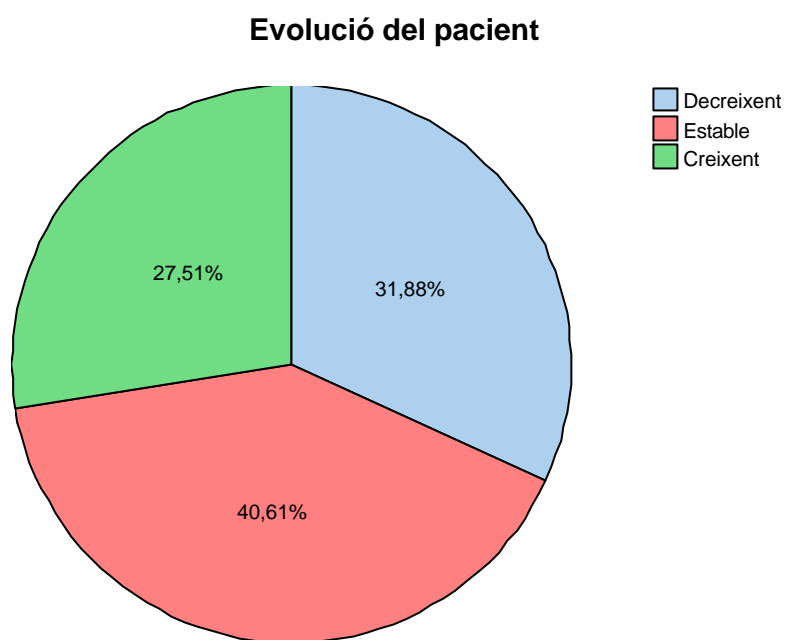


Figura 9: Evolució del pacient

Evolució del pacient	
Decreixent	73
Estable	93
Creixent	63

A l'eliminar individus amb una sola visita, hem eliminat aquelles observacions on l'evolució era estable, hem passat de tenir un 72,3% (veure **figura 4**) dels pacients amb evolució estable a un 40,6%.

Astigmatisme

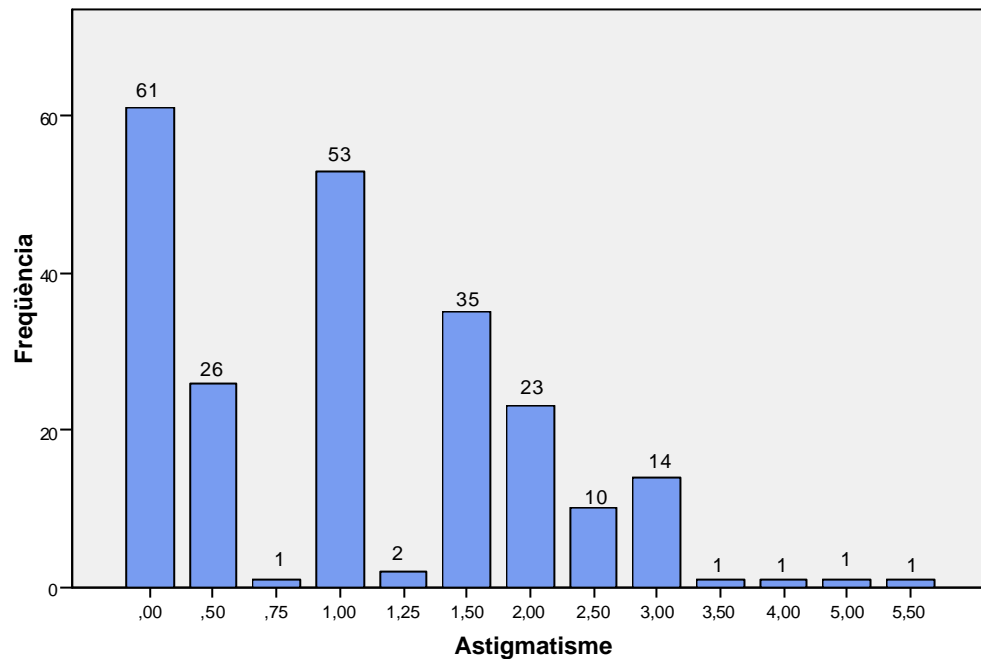


Figura 10: Gràfic d'astigmatisme

Edat a la primera visita		Equivalent esfèric	
Min.	1	Min.	-14
1st. Qu:	15	1st. Qu:	-1
Median	34	Median	1,5
Mean	48,57	Mean	0,40993
3rd.Qu:	72,5	3rd.Qu:	3
Max.	171	Max.	8

Període de seguiment		Astigmatisme	
Min.	18	Min.	0
1st. Qu:	64	1st. Qu:	0
Median	89	Median	1
Mean	90,51	Mean	1,1037
3rd.Qu:	114	3rd.Qu:	1,5
Max.	160	Max.	5,5

Els pacients amb més de dues visites, tenien una edat compresa entre 1 mes i 171 mesos.

Podem veure que la distribució dels valors de l'equivalent esfèric es troba lleugerament desplaçada cap a la dreta, ja que tant la mitjana com la mediana tenen valors positius, això ens indica que tenim més pacients amb hipermetropia que amb miopia.

El període de seguiment no és el mateix per a cada pacient. La mitjana d'aquest període és de 90 mesos, però ens trobem amb pacient que s'han estat visitant únicament durant 18 mesos i altres que han estat visitant-se durant més temps. El temps màxim de seguiment ha estat de 160 mesos.

El 25% dels pacients amb més de dues visites, no tenen astigmatisme, però tenim uns pacients que tenen un grau d'astigmatisme elevat (superior al 4.0).

L'objectiu d'aquest estudi és observar com evoluciona l'equivalent esfèric al llarg del temps, tenint en compte les dades agrupades per nen i ull.

4.3. Model Lineal Simple

Per ajustar l'equivalent esfèric en funció de l'edat, la primera idea que podem tenir és ajustar un model lineal simple:

$$y_{ijk} = x_{ijk} \mathbf{b} + \mathbf{e}_{ijk}$$

On:

- y és la variable resposta que ens indica l'equivalent esfèric,
- x és la covariable principal que ens indica l'edat del pacient en mesos,
- \mathbf{e} és l'error,
- i ens indica el pacient,
- j ens indica l'ull i
- k ens indica l'observació.

Per aplicar aquest model, necessitem que les observacions segueixin una mostra aleatòria simple, i en el nostre cas no és així. Per poder garantir que tenim una mostra aleatòria simple, només podem tenir una única observació per a cada nen i d'un ull en concret.

El model obtingut, ajustant el model lineal simple, **no és un model correcte**, ja que la principal hipòtesi no es compleix. A més a més, les nostres dades tenen variables d'agrupació i aquest model no les té en compte.

A continuació podem veure els resultats obtinguts a l'ajustar el model lineal simple:

```
Call:
lm(formula = ee ~ mes, data = dades)

Coefficients:
(Intercept)      mes
    1.82167    -0.01580

    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.0158  0.4436  0.9029  0.9029  1.3620  1.8220

Analysis of Variance Table

Response: ee
          Df Sum Sq Mean Sq F value    Pr(>F)
mes         1   1475    1475   79.129 < 2.2e-16 ***
Residuals 1844   34384      19
---
          1.82167    -0.01580

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model lineal simple amb dades agrupades

Agrupació a un nivell

Per cada nen tenim més d'una observació, per tant, podem ajustar un model lineal tenint en compte que les dades estan agrupades per nen.

Aquest model **no és correcte**, ja que només tenim en compte una de les agrupacions. Les dades tenen dos nivells d'agrupació, nen i ull, al no utilitzar l'agrupació de l'ull estem ometent informació. El model matricial és:

$$Y_i = X_i \mathbf{b}_i + \mathbf{e}_i$$

Per a cada nen tenim una mitjana i una pendent diferents.

```
Call:
lm Model: ee ~ mes | id
Data: dades1
```

Coefficients:

```
              (Intercept)              mes
372 -15.31439556  7.582230e-03
94  -12.73253414 -6.748174e-04
14   -8.05431971 -2.496120e-02
132  -9.63978495 -3.494624e-02
65   -11.26827318  1.433303e-02
...
457   6.56216472 -3.707794e-03
760   5.43287037  4.629630e-03
674   7.69475863 -6.412846e-03
42    7.49876431  6.272289e-03
358   5.45603225  3.414533e-02

Degrees of freedom: 1846 total; 1388 residual

Residual standard error: 1.197954
```

Agrupació a dos nivell

Ajustem un model lineal agrupant per nen i ull, per tant tindrem un model per a cada nen i per a cada ull. El model de forma matricial és:

$$Y_{ij} = X_{ij} \mathbf{b}_{ij} + \mathbf{e}_{ij}$$

Call:

```
Model: ee ~ mes | id/eye
Data: dades2
```

Coefficients:

```
              (Intercept)              mes
1/0   1.453720e+00 -1.149647e-02
1/1   1.363420e+00 -1.088344e-02
4/0   3.952584e+00 -5.459477e-03
4/1   2.492982e+00 -8.259660e-03
8/0  -5.863299e-01 -4.439588e-03
8/1   9.642343e-01 -1.721896e-02
...
760/1  3.356481e+00  1.851852e-02
761/0 -4.870945e+00  4.887378e-03
```

```

761/1 -2.375761e+00 -3.718657e-04
762/0  2.661811e+00  1.824344e-02
762/1  2.970003e+00  2.112778e-02

```

Degrees of freedom: 1846 total; 930 residual

Residual standard error: 0.6033024

	(Intercept)	mes	visit	Nvisits	ini	period	ee	ast	sexe	enf	evol
1/0	0.657441308	-8.290442e-03	3.5	6	22	115	0.541666667	0.750000000	0	0	-1
1/1	0.567140670	-7.677408e-03	3.5	6	22	115	0.500000000	1.666666667	0	0	-1
4/0	3.156305126	-2.253445e-03	3.5	6	88	160	3.020833333	1.291666667	0	1	0
4/1	1.696703051	-5.053629e-03	3.5	6	88	160	1.083333333	0.416666667	0	0	-1
8/0	-1.382608820	-1.233557e-03	2.0	3	80	77	-1.125000000	0.583333333	1	-1	0
8/1	0.167955464	-1.401293e-02	2.0	3	80	77	-1.125000000	0.750000000	1	0	-1
9/0	-0.270777297	2.146168e-02	3.0	5	19	105	1.975000000	0.250000000	0	0	1
...											
759/0	1.665177337	1.288207e-02	2.5	4	59	111	3.625000000	0.625000000	1	1	1
759/1	2.142314798	1.099319e-02	2.5	4	59	111	3.875000000	0.625000000	1	1	1
760/0	6.712980379	-6.053228e-03	2.0	3	64	54	6.666666667	2.000000000	1	1	0
760/1	2.560202602	2.172455e-02	2.0	3	64	54	5.041666667	1.750000000	1	1	1
761/0	-5.667223773	8.093409e-03	2.0	3	62	97	-4.333333333	1.000000000	1	-1	0
761/1	-3.172040319	2.834166e-03	2.0	3	62	97	-2.416666667	1.500000000	1	-1	0
762/0	1.865532484	2.144947e-02	2.5	4	28	76	3.875000000	1.625000000	1	1	1
762/1	2.173724004	2.433381e-02	2.5	4	28	76	4.375000000	2.000000000	1	1	1

Aquest model **és correcte**, a més a més, és el model que ens donarà un error residual més petit (**(0.6033)²**). Al ser un model saturat i sense covariables, fa que no el puguem utilitzar per fer inferència. Per tant, no podem fer prediccions per nos nens perquè no té estructura poblacional i és un model que té molts paràmetres.

Si ens fixem en els coeficients obtinguts per a cada nen són molt diferents entre els nens, pero són molt similars entre els obtinguts per l'ull dret i l'ull esquerre de cada pacient.

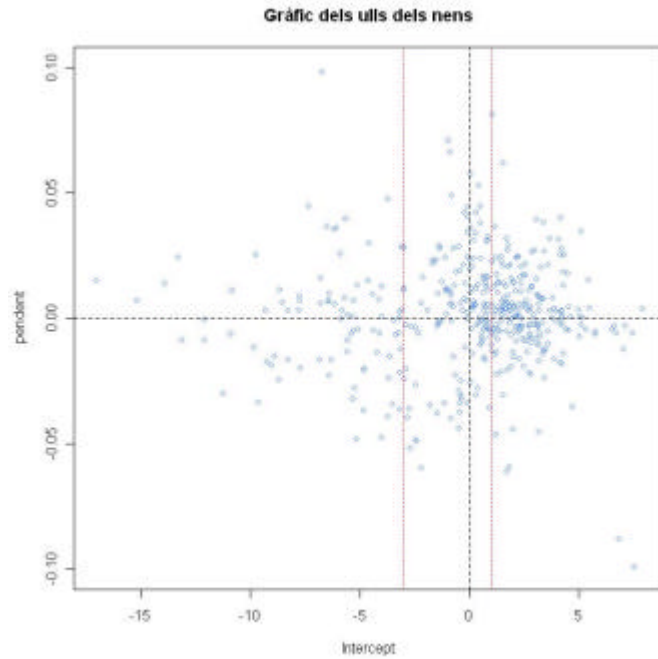


Figura 11: Gràfic dels ulls dels nens

Els nens que tenen una pendent positiva i un intercept positiu, són nens que van començar l'estudi amb una hipermetropia que els ha anat augmentant a mesura que avançava l'estudi. Per contra, els nens amb els dos valors negatius, són nens que al inici tenien miopia i que els ha anat augmentant. Els nens que trobem amb una pendent pròxima a zero són nens que s'han mantingut estables al llarg del estudi.

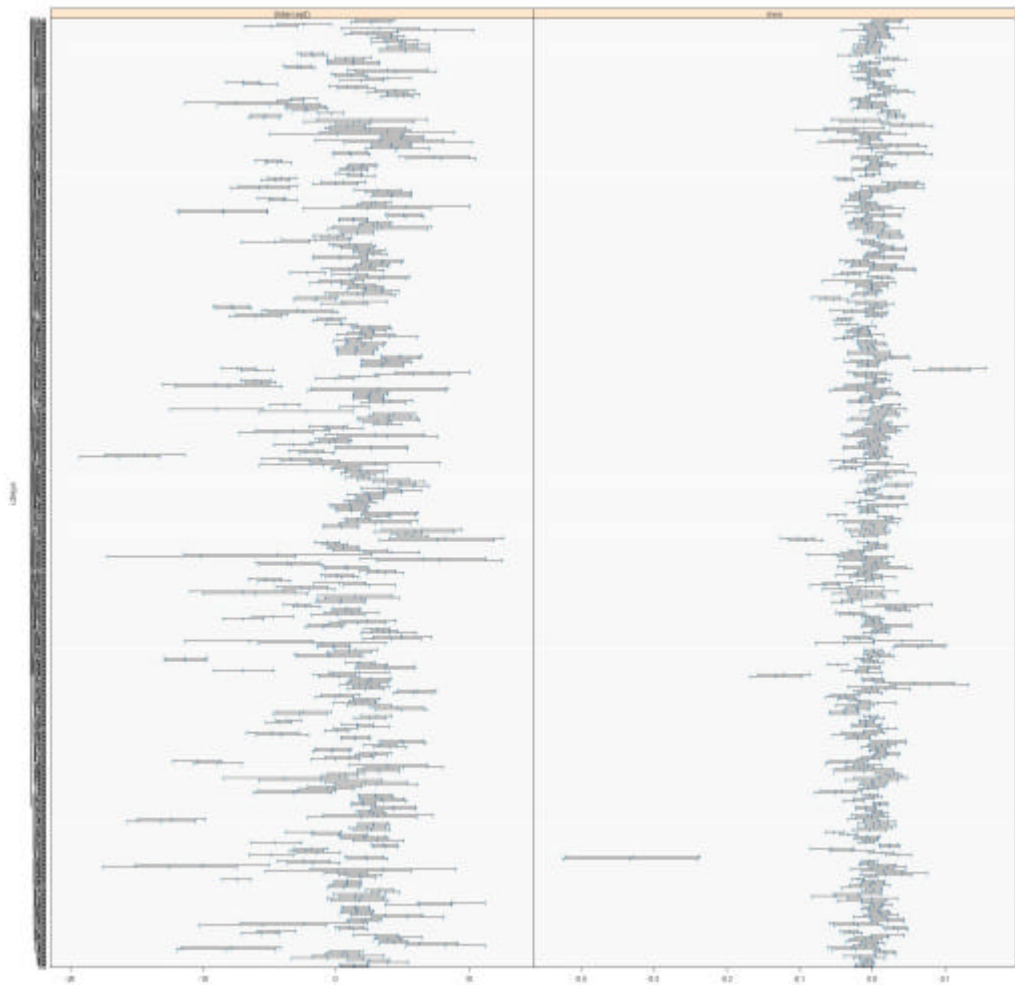


Figura 12: Gràfic dels intervals de confiança pels paràmetres del model

Que els intervals de confiança no es distribueixin de manera constant, ens indica que necessitarem incloure en el nostre model la part d'efectes aleatoris.

4.4. Models lineals mixtos

Model 1

El primer model mixt que podem ajustar consisteix en una millora de la regressió lineal simple considerant dos nivells d'agrupació, nen i ull.

$$Y_{ij} = \underbrace{X_{ij}\mathbf{b}}_{\text{Part fixa}} + \underbrace{Z_{i,j}b_i + Z_{ij}b_{ij}}_{\text{Part aleatòria}} + \mathbf{e}_{ij}$$

Els paràmetres β són els valors promig de la població, la b_i és la variància associada al nen i-èsim i la b_{ij} és la variància deguda a la mesura realitzada en l'ull j del pacient i.

La matriu de disseny és:

$$\begin{pmatrix} y_{1ij} \\ y_{2ij} \\ \vdots \\ y_{nij} \end{pmatrix} = \begin{pmatrix} 1 & x_{1ij} \\ 1 & x_{2ij} \\ \vdots & \vdots \\ 1 & x_{nij} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} + \begin{pmatrix} 1 & x_{1ij} \\ 1 & x_{2ij} \\ \vdots & \vdots \\ 1 & x_{nij} \end{pmatrix} \begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} + \begin{pmatrix} 1 & x_{1ij} \\ 1 & x_{2ij} \\ \vdots & \vdots \\ 1 & x_{nij} \end{pmatrix} \begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} + \begin{pmatrix} e_{1ij} \\ e_{2ij} \\ \vdots \\ e_{nij} \end{pmatrix}$$

Hipòtesis de normalitat del primer nivell (Nivell nen):

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)}) \quad V^{(1)} = \begin{pmatrix} s_1^{(1)} & 0 \\ 0 & s_2^{(1)} \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & r_{12}^{(1)} \\ r_{12}^{(1)} & 1 \end{pmatrix}$$

Hipòtesis de normalitat del segon nivell (Nivell ull):

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)}) \quad V^{(2)} = \begin{pmatrix} s_1^{(2)} & 0 \\ 0 & s_2^{(2)} \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & r_{12}^{(2)} \\ r_{12}^{(2)} & 1 \end{pmatrix}$$

Hipòtesis de normalitat dels residus:

$$e_{ijk} \sim N(0, s^2 \Lambda)$$

Podem reescriure el model com:

$$ee_{ijk} = (b_0 + b_{0i} + b_{0ij}) + (b_1 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$$

En el model ajustat no hem donat cap estructura als efectes aleatoris.

Linear mixed-effects model fit by REML

Random effects:

Formula: ~mes | id

Structure: General positive-definite

StdDev Corr

(Intercept) 3.60464868 (Intr)

mes 0.02081405 0.101

```

Formula: ~mes | eye %in% id

Structure: General positive-definite

          StdDev      Corr

(Intercept) 0.663096044 (Intr)
mes          0.004059282 1
Residual     0.607302387

Fixed effects: ee ~ mes

          Value      Std.Error    DF    t-value  p-value
(Intercept) 0.7830255 0.24497645 1387   3.196330  0.0014
mes         -0.0025461 0.00147554 1387  -1.725509  0.0847
Correlation:
      (Intr)
mes 0.051

Standardized Within-Group Residuals:

          Min          Q1          Med          Q3          Max
-4.35724255 -0.43062845  0.00543532  0.47381737  3.82417769

Number of Observations: 1846

Number of Groups:

      id eye %in% id
      229      458

```

El model obtingut és:

$$ee_{ijk} = (0.7830 + b_{0i} + b_{0ij}) + (-0.0025 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$$

Nivell nen:

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)}C^{(1)}V^{(1)}) \quad V^{(1)} = \begin{pmatrix} 3.6046 & 0 \\ 0 & 0.0208 \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & 0.101 \\ 0.101 & 1 \end{pmatrix}$$

Nivell ull:

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)}C^{(2)}V^{(2)}) \quad V^{(2)} = \begin{pmatrix} 0.6631 & 0 \\ 0 & 0.0041 \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Residus:

$$e_{ijk} \sim N(0, (0.6073)^2 I)$$

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	0.302461549	0.783025609	1.2635896681
mes	-0.005440599	-0.002546063	0.0003484737

```
attr(,"label")
[1] "Fixed effects:"
```

Random Effects:

```
Level: id
```

	lower	est.	upper
sd((Intercept))	3.26368682	3.60464594	3.98122524
sd(mes)	0.01842093	0.02081405	0.02351807
cor((Intercept),mes)	-0.04287745	0.10087896	0.24054290

```
Level: eye
```

	lower	est.	upper
sd((Intercept))	0.550712318	0.663100429	0.7984244
sd(mes)	0.003214635	0.004059298	0.0051259
cor((Intercept),mes)	-1.000000000	0.999999648	1.0000000

Within-group standard error:

lower	est.	upper
0.5822310	0.6073018	0.6334522

Si els intervals de confiança pels paràmetres inclouen el zero, el **test de Wald** ens indica que el valor estimat no és significatiu, és a dir, no aporta cap informació al model. En aquest cas, podem tornar a definir el model eliminant els coeficients que no ens aporten noves informacions.

En el nostre cas, tenim tres paràmetres que no són significatius, la pendent del efectes fixes \mathbf{b}_0 , i les correlacions dels efectes aleatoris $\mathbf{r}_{12}^{(1)}$ i $\mathbf{r}_{12}^{(2)}$. Simplificarem el model ajustat eliminant un dels paràmetres no significatius. Començarem traient un dels paràmetres dels efectes aleatoris ja que el paràmetre de la part fixa és de la covariable principal del model.

Simplificarem el model donant estructura diagonal a la matriu de variàncies i covariàncies del primer nivell, ja que el coeficient estimat està molt a prop del zero, en el cas del segon nivell el coeficient s'aproxima a 1.

Model 2

Per eliminar la correlació de primer nivell dels efectes aleatoris forcem que la matriu de correlacions sigui diagonal, és a dir, consideren que la correlació és zero.

Linear mixed-effects model fit by REML

Random effects:

Formula: ~mes | id

Structure: Diagonal

(Intercept) mes

StdDev: 3.625067 0.02097457

Formula: ~mes | eye %in% id

Structure: General positive-definite

StdDev Corr

(Intercept) 0.664245066 (Intr)

mes 0.004066294 1

Residual 0.606160032

Fixed effects: ee ~ mes

	Value	Std.Error	DF	t-value	p-value
(Intercept)	0.7843746	0.24633357	1387	3.184197	0.0015
mes	-0.0025597	0.00148578	1387	-1.722793	0.0851

Correlation:

(Intr)

mes -0.04

Standardized Within-Group Residuals:

	Min	Q1	Med	Q3	Max
	-4.362231242	-0.426870823	0.006069872	0.473340873	3.830731113

Number of Observations: 1846

Number of Groups:

```
id eye %in% id
229      458
```

El model obtingut és:

$$ee_{ijk} = (0.7844 + b_{0i} + b_{0ij}) + (-0.0026 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$$

Nivell nen:

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)}) \quad V^{(1)} = \begin{pmatrix} 3.6251 & 0 \\ 0 & 0.0210 \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Nivell ull:

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)}) \quad V^{(2)} = \begin{pmatrix} 0.6642 & 0 \\ 0 & 0.0041 \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Residus:

$$e_{ijk} \sim N(0, (0.6062)^2 I)$$

La mitjana poblacional ha augmentat en comparació amb el model anterior (de 0.7830 a 0.7844). La correlació de segon nivell és una mica grollera, perquè assumim que els dos ulls es comporten de manera similar.

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	Upper
(Intercept)	0.301147977	0.784374592	1.2676012065
mes	-0.005474309	-0.002559691	0.0003549273

```
attr(,"label")
[1] "Fixed effects:"
```

Random Effects:

Level: id	lower	est.	upper
sd((Intercept))	3.28321053	3.62506722	4.00251895
sd(mes)	0.01857351	0.02097457	0.02368603

Level: eye			
	lower	est.	upper
sd((Intercept))	0.549229858	0.664245066	0.803345815
sd(mes)	0.003217889	0.004066294	0.005138383
cor((Intercept),mes)	-1.000000000	0.999999710	1.000000000
Within-group standard error:			
	lower	est.	upper
	0.5812762	0.6061600	0.6321092

Si ens fixem en els intervals de confiança el paràmetre b_0 de la part fixa i el paràmetre $r_{12}^{(2)}$ de la part aleatòria continuen sense ser significatius. No ens interessa eliminar el coeficient de la part fixa perquè és el paràmetre de la covariable principal, per tant, mirarem de simplificar el model eliminant el paràmetre de la part aleatòria. A partir del Test de Wald podem decidir eliminar el paràmetre però és una dada amb restricció a la frontera i el interval de confiança ens dóna tot el rang de possibles valors. Per tant, aplicarem el test de la raó de versemblança. Per fer-ho, ajustarem un nou model sense aquesta variable i comprovarem si el model ha millorat.

Model 3

Ajustem un nou model eliminant la correlació del segon nivell dels efectes aleatoris, per fer-ho, forcem la matriu de correlacions perquè sigui diagonal, és a dir, consideren que la correlació és zero.

Linear mixed-effects model fit by REML	
Random effects:	
Formula:	~mes id
Structure:	Diagonal
	(Intercept) mes
StdDev:	3.624333 0.02091237
Formula:	~mes eye %in% id

```

Structure: Diagonal

      (Intercept)      mes
StdDev:   0.7976423 0.006167586

Residual    0.598192
Fixed effects: ee ~ mes
              Value Std.Error   DF   t-value p-value
(Intercept)  0.7849753 0.24705135 1387   3.177377   0.0015
mes          -0.0025652 0.00149592 1387  -1.714776   0.0866

Correlation:
      (Intr)
mes -0.054

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-4.333547079 -0.417765556  0.002189271  0.454538366  3.586702278

Number of Observations: 1846
Number of Groups:
      id eye %in% id
      229      458

```

El model obtingut és:

$$ee_{ijk} = (0.7850 + b_{0i} + b_{0ij}) + (-0.0026 + b_{1i} + b_{1ij}) \dots mes_{ijk} + e_{ijk}$$

Nivell nen:

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)}) \quad V^{(1)} = \begin{pmatrix} 3.6243 & 0 \\ 0 & 0.0209 \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Nivell ull:

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)}) \quad V^{(2)} = \begin{pmatrix} 0.7976 & 0 \\ 0 & 0.0062 \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Residus:

$$e_{ijk} \sim N(0, (0.5982)^2 I)$$

Approximate 95% confidence intervals
Fixed effects:

	lower	est.	upper
(Intercept)	0.300340615	0.784975268	1.2696099210
mes	-0.005499674	-0.002565166	0.0003693425

```
attr(,"label")
[1] "Fixed effects:"
```

Random Effects:

```
Level: id
```

	lower	est.	upper
sd((Intercept))	3.28118190	3.62433285	4.00337104
sd(mes)	0.01847209	0.02091237	0.02367502

```
Level: eye
```

	lower	est.	upper
sd((Intercept))	0.678650298	0.797642268	0.937497838
sd(mes)	0.005123151	0.006167586	0.007424946

Within-group standard error:

	lower	est.	upper
	0.5728813	0.5981920	0.6246210

Com que teníem un valor amb restricció de frontera, hem fet el test de raó de versemblança per saber si el model millora quan el traïem.

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
	model.lme32	1	8	5855.012	5899.169	-2919.506		
	model.lme33	2	7	5878.103	5916.741	-2932.052	1 vs 2	25.09106 <.0001

Amb aquest test comprovem si existeixen diferències significatives entre el model amb la correlació de segon nivell i el model sense la correlació. Podem veure que no existeixen diferències significatives entre un model i l'altre, per tant ens quedarem amb aquell model que tingui l'estadístic AIC més petit, o el que és mateix, aquell que

tingui la funció log-likelihood més gran. En el nostre cas ens quedem amb el model que té més paràmetres.

El model que ens queda és:

$$ee_{ijk} = (0.7844 + b_{0i} + b_{0ij}) + (-0.0025 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$$

però hem de validar si compleix les hipòtesis de linealitat, independència i normalitat.

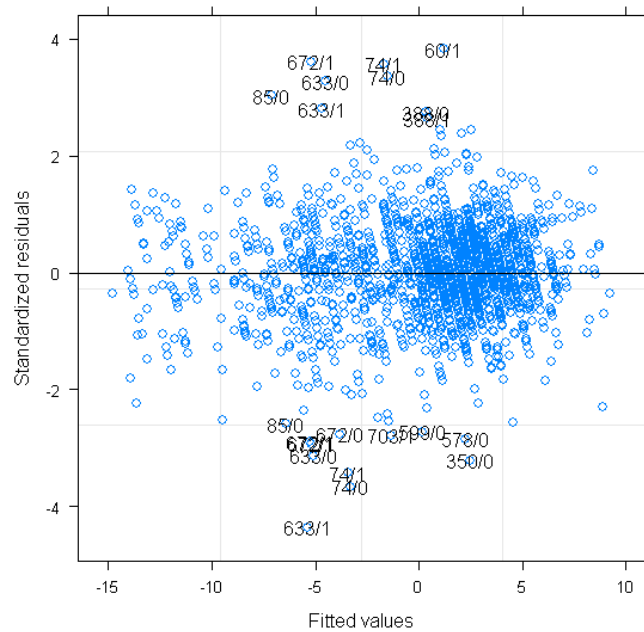


Figura 13: Gràfic dels valors estimats dels model vers els residus estandarditzats

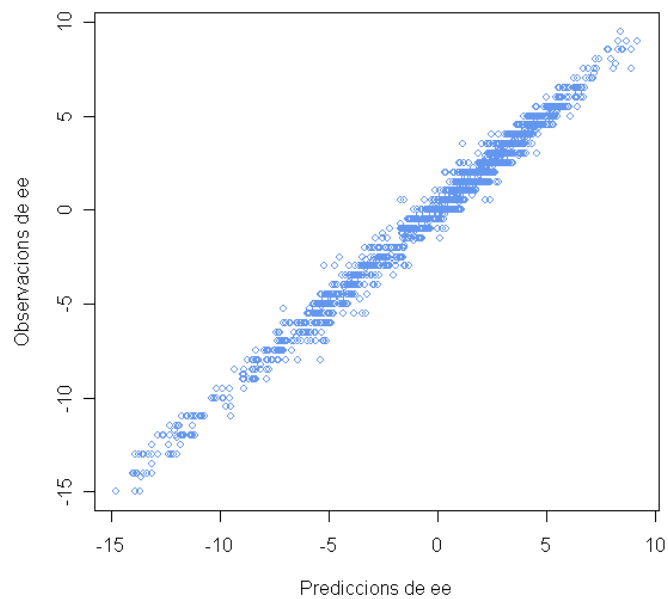


Figura 14: Gràfic dels valors predits vers les observacions del model

En el gràfic dels valors predits vers els errors estandarditzats (**figura 13**) podem veure que algunes prediccions ens donen uns errors molt grans, això pot ser degut a la presència de valor extrems o outliers en la mostra. En la **figura 14**, observem que l'error entre la predicció i l'observació no és molt gran, per tant el model lineal és bo. En aquests gràfics podem veure que les dades semblen agrupades en franges, això és perquè l'equivalent esfèric té una precisió de 0'25.

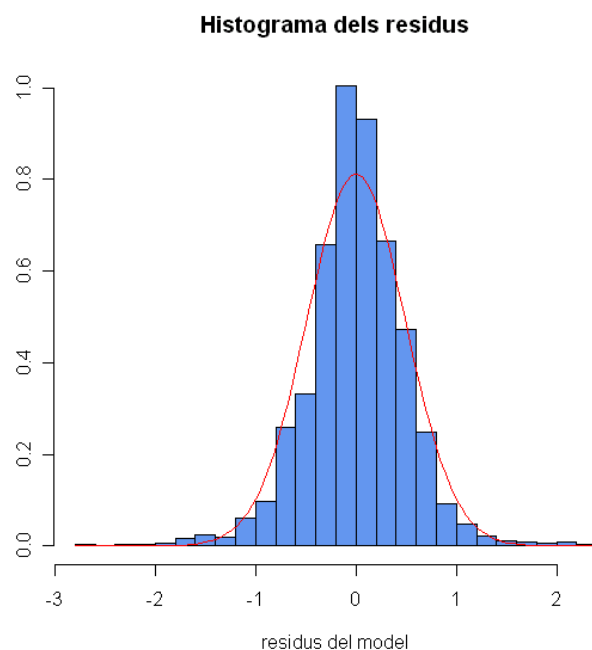


Figura 15: Histograma dels residus del model amb la corba de la normal

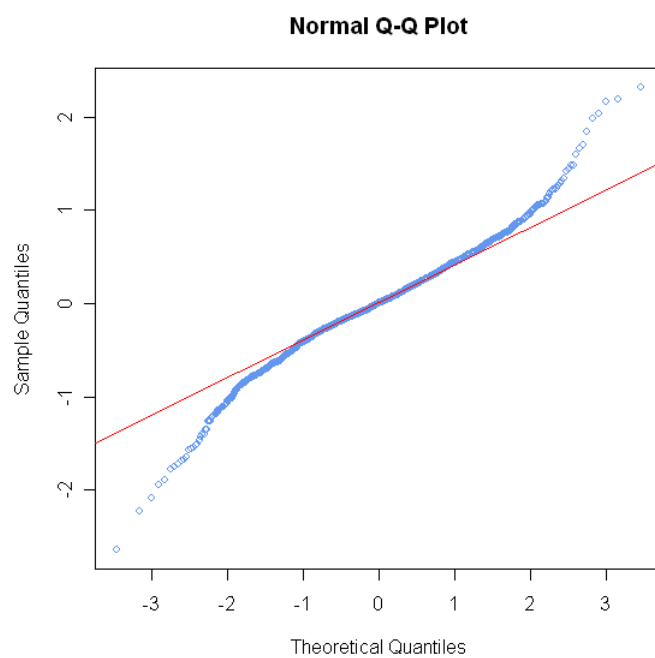


Figura 16: Qqnorm dels residus del model

Si mirem l'histograma dels residus (**figura 15**) podem dir que els errors del model segueixen una distribució normal tot i que està desplaçada una mica cap a l'esquerra. El gràfic qqnorm ens serveix per comprovar la normalitat de les dades. En aquest

gràfic, podem veure que les dades tenen unes cues pesants. Això pot ser perquè tinguem valors extrems a les dades que estem estudiant.

Per comprovar les hipòtesis dels efectes aleatoris, mirem si tenim variància constant, és a dir, homoscedasticitat.

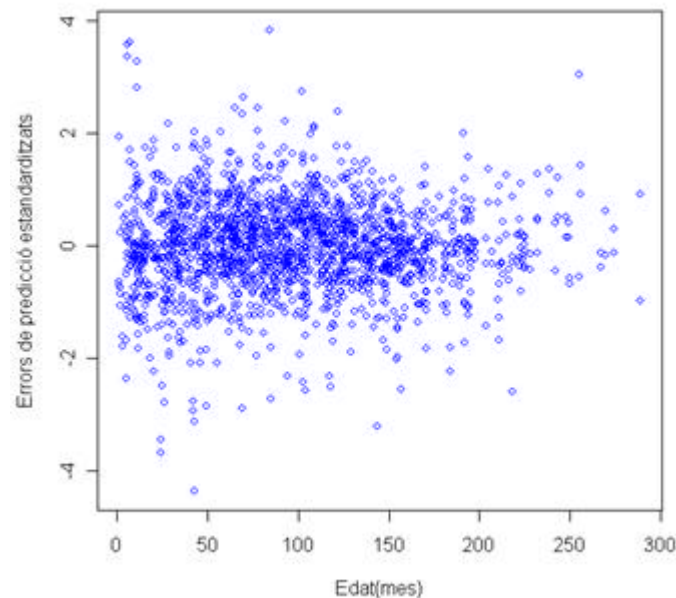


Figura 17: Scatter-plot dels errors de predicció estandarditzats vers l'edat en mesos

En la figura anterior, observem que la variància disminueix a mesura que augmenta l'edat, això és normal ja que a mesura que ens fem grans, tendim a estabilitzar l'equivalent esfèric. Per tant, la variància dels efectes aleatoris no és constant. Per aconseguir una variància constant podem aplicar una potència.

Model 4

En el model anterior, hem observat que la variància dels residus no és constant, disminueix a mesura que la covariable principal (mes) augmenta. Provem d'ajustar un nou model afegint una potència a la covariable de la variància del model.

Linear mixed-effects model fit by REML

Fixed: ee ~ mes

(Intercept) mes

0.827180663 -0.003213847

Random effects:

```
Formula: ~mes | id
Structure: Diagonal
      (Intercept)      mes
StdDev:   3.724979 0.01776182

Formula: ~mes | eye %in% id
Structure: General positive-definite
      StdDev      Corr
(Intercept) 0.740854354 (Intr)
mes          0.003573787 1

Residual    1.759128335
```

Variance function:

```
Structure: Power of variance covariate
Formula: ~mes
Parameter estimates:
      power
-0.2536237

Number of Observations: 1846

Number of Groups:
      id eye %in% id
      229      458
```

El model obtingut és:

$$ee_{ijk} = (0.8272 + b_{0i} + b_{0ij}) + (-0.0032 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$$

Nivell nen:

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)}C^{(1)}V^{(1)}) \quad V^{(1)} = \begin{pmatrix} 3.7250 & 0 \\ 0 & 0.0178 \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Nivell ull:

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)}C^{(2)}V^{(2)}) \quad V^{(2)} = \begin{pmatrix} 0.7408 & 0 \\ 0 & 0.0036 \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Residus:

$$e_{ijk} \sim N\left(0, (1.7591)^2 \text{mes}^{-0.2536}\right)$$

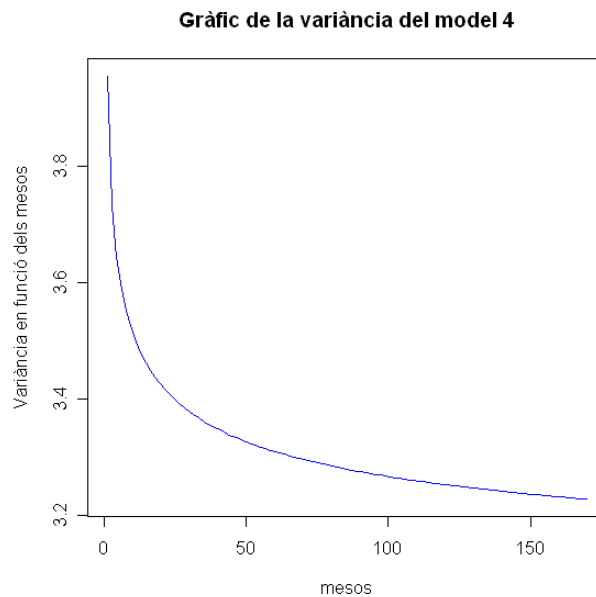


Figura 18: Gràfic de la variància dels residus en funció de l'edat en el model 4

A l'aplicar una funció a la variància, el model es complica i el test de Wald no convergeix, per tant, per saber si el model és millor ara, utilitzarem el test de la raó de versemblança.

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
	model.lme32	1 8	5855.012	5899.169	-2919.506			
	model.lme34	2 9	5778.376	5828.053	-2880.188	1 vs 2	78.63599	<.0001

En el nostre cas, el model no té diferències significatives, per tant, ens quedem amb aquell que té l'estadístic AIC menor, tot i que, introduïm un paràmetre més al model.

Model 5

Podem intentar millorar el model anterior aplicant una altra estructura a la matriu de variàncies i covariàncies dels residus. La funció `conspower`, a més d'aplicar una potència a la covariable principal, li suma una constant.

Linear mixed-effects model fit by REML

Fixed: ee ~ mes

(Intercept) mes

0.835614054 -0.003337007

Random effects:

Formula: ~mes | id

Structure: Diagonal

(Intercept) mes

StdDev: 3.727371 0.01759443

Formula: ~mes | eye %in% id

Structure: General positive-definite

StdDev Corr

(Intercept) 0.751536518 (Intr)

mes 0.003512811 0.996

Residual 4.232777113

Variance function:

Structure: Constant plus power of variance covariate

Formula: ~mes

Parameter estimates:

const power

0.1043517 -0.8294338

Number of Observations: 1846**Number of Groups:**

id eye %in% id

229 458

El model obtingut és:

$$ee_{ijk} = (0.8356 + b_{0i} + b_{0ij}) + (-0.0033 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$$

Nivell nen:

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)}) \quad V^{(1)} = \begin{pmatrix} 3.7274 & 0 \\ 0 & 0.0176 \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Nivell ull:

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)}) \quad V^{(2)} = \begin{pmatrix} 0.7515 & 0 \\ 0 & 0.0035 \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & 0.996 \\ 0.996 & 1 \end{pmatrix}$$

Residus:

$$e_{ijk} \sim N(0, (4.2328)^2 (0.1043 + mes^{-0.8294}))$$

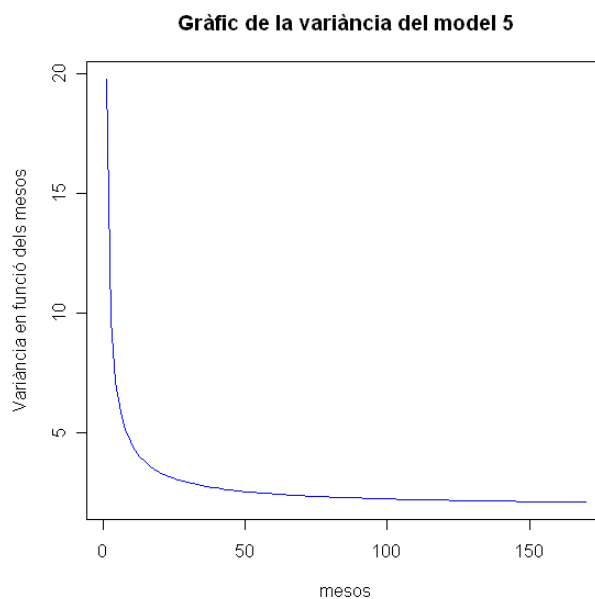


Figura 19: Gràfic de la variància dels residus en funció de l'edat en el model 5

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	0.338844542	0.835614054	1.3323835672
mes	-0.005849304	-0.003337007	-0.0008247104

```
attr(,"label")
[1] "Fixed effects:"
```

Random Effects:

```
Level: id
```

	lower	est.	upper
sd((Intercept))	3.37868938	3.72737081	4.11203624
sd(mes)	0.01558241	0.01759443	0.01986625

Level: eye			
	lower	est.	upper
sd((Intercept))	0.637282574	0.751536518	0.886274254
sd(mes)	0.002668298	0.003512811	0.004624611
cor((Intercept),mes)	-1.000000000	0.995535367	1.000000000
Variance function:			
	lower	est.	upper
const	0.05654012	0.1043517	0.1925938
power	-1.14482841	-0.8294338	-0.5140392
attr(,"label")			
[1] "Variance function:"			
Within-group standard error:			
	lower	est.	upper
	2.059031	4.232777	8.701375

Pel test de Wald, podem comprovar que els paràmetres estimats per la variància són significativament diferents de zero.

Pel test de Likelihood Ratio, comprovarem si el model aplicant una constant i una potència a la covariable **mes**, ens ajusta millor que el model aplicant únicament una potència.

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
	model.lme34	1	9	5778.376	5828.053	-2880.188		
	model.lme35	2	10	5766.608	5821.805	-2873.304	1 vs 2	13.76753
								2e-04

El test de Likelihood Ratio no ens mostra diferències significatives entre els models, per tant, ens quedem aquell model que ens ajusta millor les dades. El model que té un valor del log-likelihood major és el model amb estructura de la variància amb la funció conspower.

Comprovem si aquest model compleix les hipòtesis de linealitat, independència i normalitat.

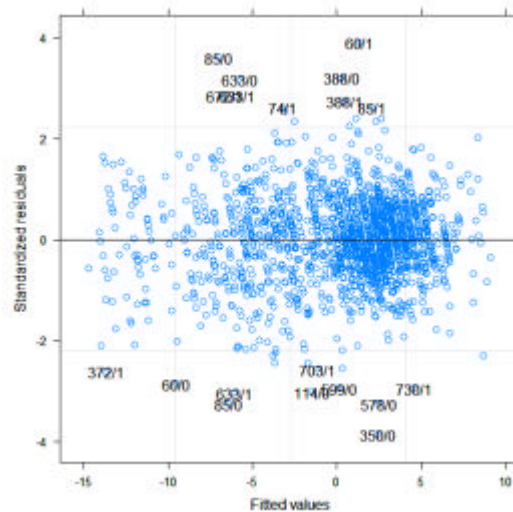


Figura 20: Gràfic dels valors estimats dels model vers els residus estandarditzats

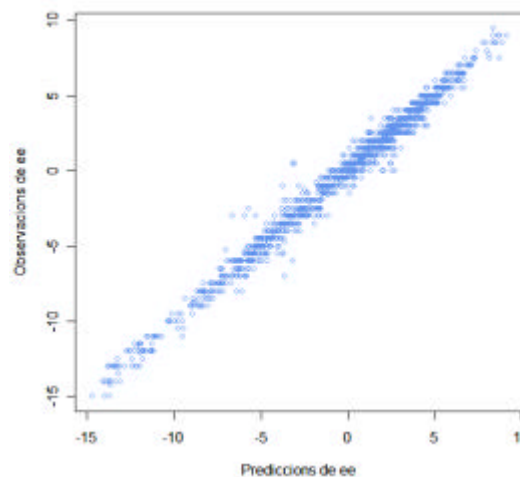


Figura 21: Gràfic dels valors predits vers les observacions del model

En la **figura 20**, observem que alguns valor ajustats ens donen uns errors molt grans, això pot ser degut a valors extrems o outliers. Per millorar el model, podem eliminar els individus que ens donen aquests errors.

En el gràfic dels valors observats amb les prediccions (**figura 21**) veiem que les hipòtesis de linealitat es compleixen.

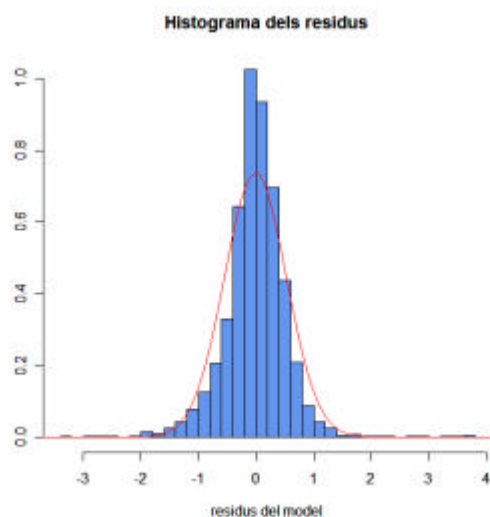


Figura 22: Histograma dels residus del model amb la corba de la normal

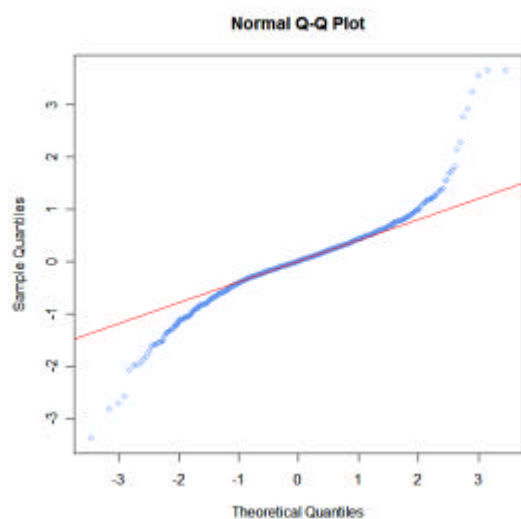


Figura 23: Qqnorm dels residus del model

Per les **figures 22 i 23**, veiem que les hipòtesis de normalitat es compleixen. En la **figura 23** observem que tenim cues pesants, això pot ser degut als valors extrems i es pot millorar si eliminem aquests valors per modelitzar les dades.

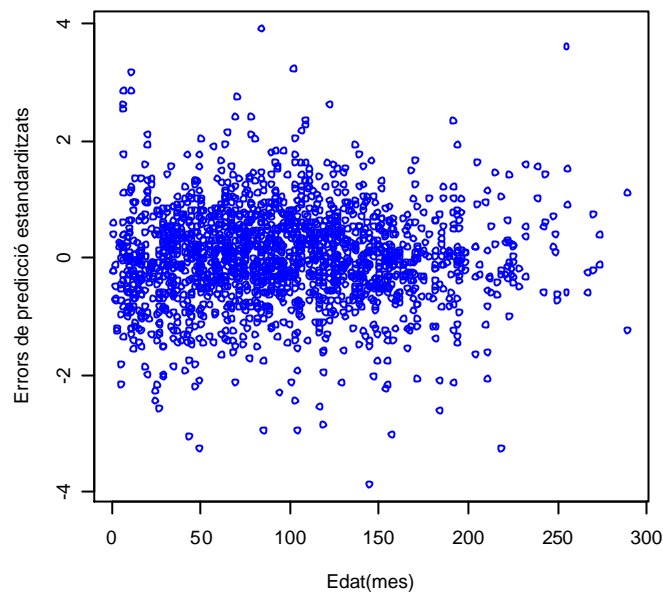


Figura 24: Scatter-plot dels errors de predicció estandarditzats vers l'edat en mesos

En el gràfic anterior veiem que la variància és més constant que en la **figura 17**, podem dir que és homoscedàstica.

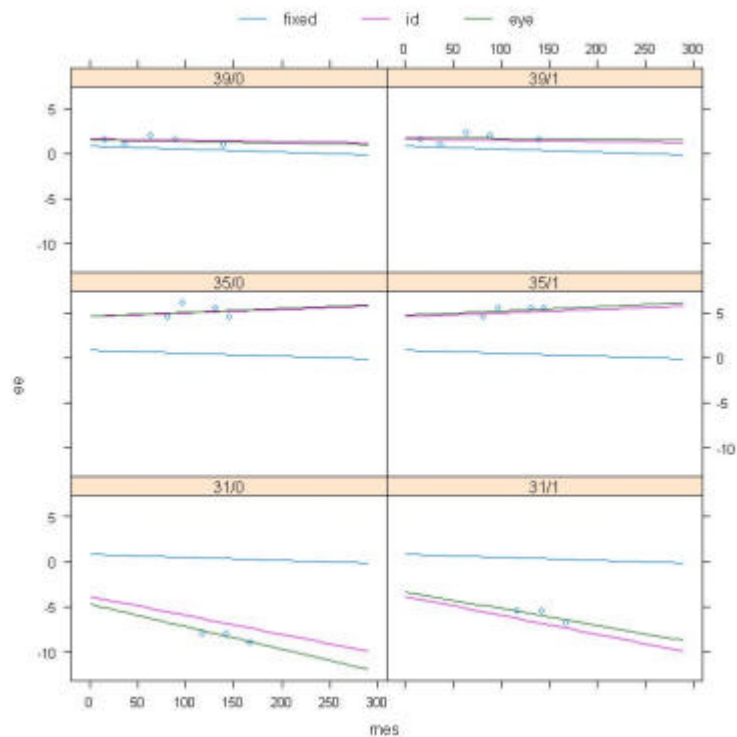


Figura 25: Modelització per 3 individus representatius de la mostra

En aquest gràfic podem veure la representació de sis ulls (tres nens). La línia de color blau representa la tendència poblacional, en color violeta podem veure la tendència del nen, i en color verd la tendència de l'ull del nen. Aquesta última línia veiem que s'ajusta molt bé amb les observacions d'aquest ull.

En el gràfic següent podem veure aquest mateix ajust del model pels nens que veiem com a valors atípics a la **figura 20**.

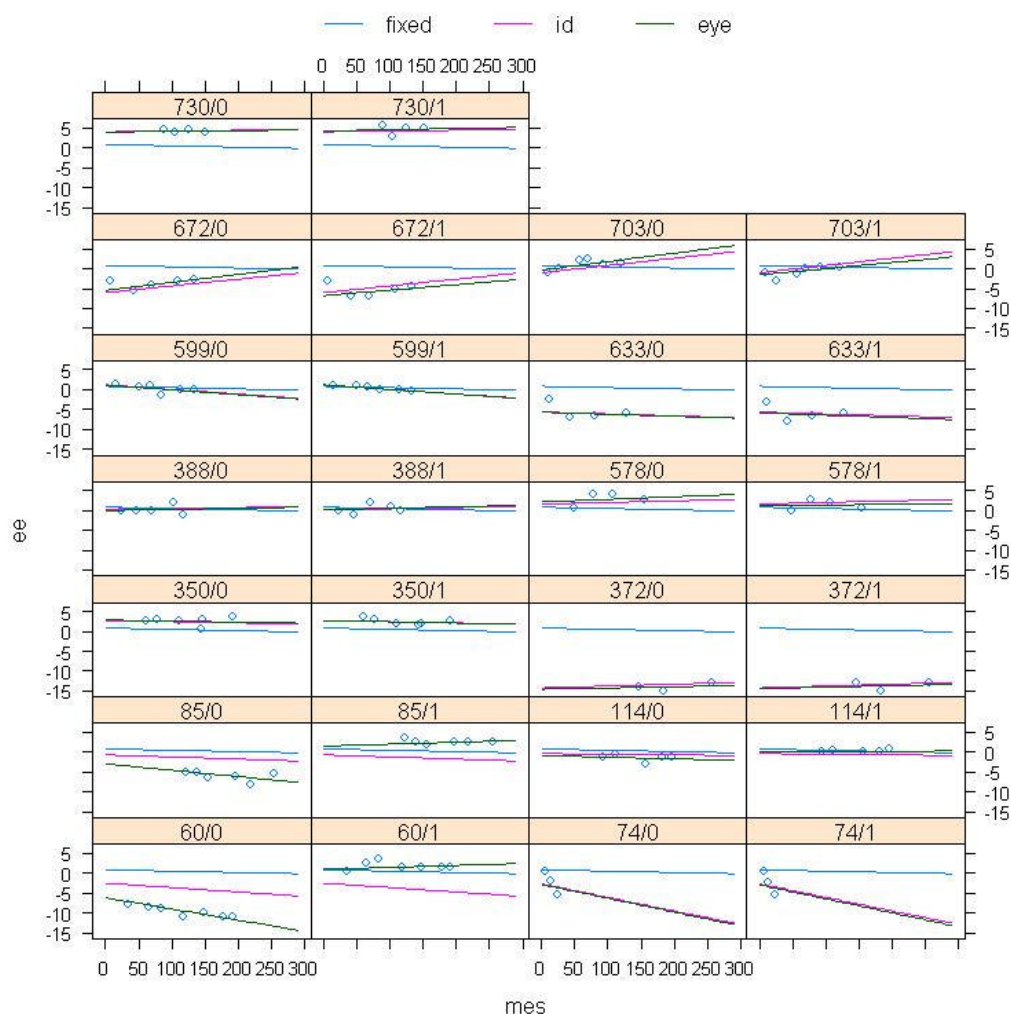


Figura 26: Modelització pels individus atípics de la mostra

Els atípics poden venir donats per 3 casos:

- Hi ha un error en algun dels mesurament de l'equivalent esfèric, com per exemple, en el cas del nen 350.

- Hi ha un comportament anòmal en l'evolució d'algun ull del nen, ja sigui per una operació o per un agreujament molt ràpid de la malaltia, com poden ser els casos 60 i 74, respectivament.
- Hi ha un comportament diferent en els dos ulls, com per exemple el cas 85.

Podem trobar aquest mateix gràfic per tots els nens de la mostra a l'**annex 7.3**.

4.5. Models lineals mixtos sense atípics

Si ens mirem la **figura 20** veurem que hi ha uns valors clarament atípics. Aquests valors són els corresponents als individus:

60	114	388	633	730
74	350	578	672	
85	372	599	703	

Com que aquests pacients tenien una influència molt gran, els hem eliminat de les nostres dades i hem tornat a ajustar el model.

Linear mixed-effects model fit by REML

Fixed: ee ~ mes

(Intercept) mes

0.982327084 -0.003073223

Random effects:

Formula: ~mes | id

Structure: Diagonal

(Intercept) mes

StdDev: **3.621043 0.01879392**

Formula: ~mes | eye %in% id

Structure: General positive-definite

StdDev Corr

(Intercept) **0.399916039** (Intr)

mes **0.004991084 0.993**

Residual **1.703969292**

Variance function:

Structure: Constant plus power of variance covariate

Formula: ~mes

Parameter estimates:

const	power
0.2094570	-0.5694911

Number of Observations: 1718

Number of Groups:

id	eye	%in%	id
216			432

El model obtingut és:

$$ee_{ijk} = (0.9823 + b_{0i} + b_{0ij}) + (-0.0031 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$$

Nivell nen:

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)}C^{(1)}V^{(1)}) \quad V^{(1)} = \begin{pmatrix} 3.6210 & 0 \\ 0 & 0.0188 \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Nivell ull:

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)}C^{(2)}V^{(2)}) \quad V^{(2)} = \begin{pmatrix} 0.3999 & 0 \\ 0 & 0.0050 \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & 0.993 \\ 0.993 & 1 \end{pmatrix}$$

Residus:

$$e_{ijk} \sim N(0, (1.7040)^2(0.2094 + mes^{-0.5695}))$$

Approximate 95% confidence intervals**Fixed effects:**

	lower	est.	upper
(Intercept)	0.490041908	0.982327084	1.4746122593
mes	-0.005787218	-0.003073223	-0.0003592279

```
attr(,"label")  
[1] "Fixed effects:"
```

Random Effects:

```

Level: id

              lower      est.      upper
sd((Intercept)) 3.2790785 3.62104287 3.99866954
sd(mes)         0.0166606 0.01879392 0.02120041

Level: eye

              lower      est.      upper
sd((Intercept)) 0.297987308 0.399916039 0.536710235
sd(mes)         0.004036237 0.004991084 0.006171819
cor((Intercept),mes) -0.999999986 0.993241637 1.000000000

```

Variance function:

```

              lower      est.      upper
const 0.1254573 0.2094570 0.3496986
power -0.8946954 -0.5694911 -0.2442869
attr(,"label")
[1] "Variance function:"

```

Within-group standard error:

```

              lower      est.      upper
0.9258403 1.7039693 3.1360822

```

Pel test de Wald, podem comprovar que els paràmetres estimats són significativament diferents de zero. Anem a fer les comprovacions del nostre model.

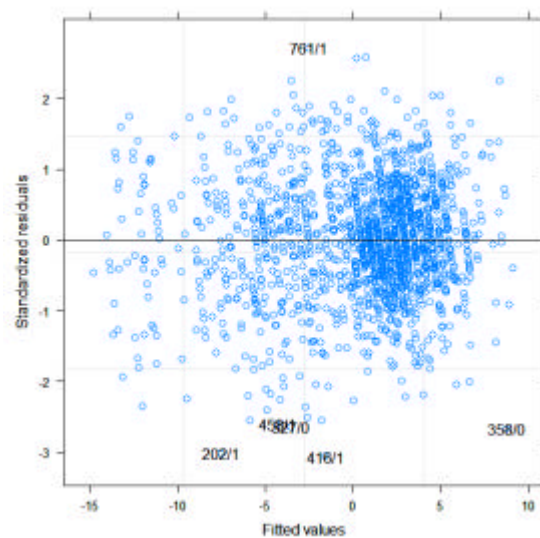


Figura 27: Gràfic dels valors estimats dels model vers els residus estandarditzats

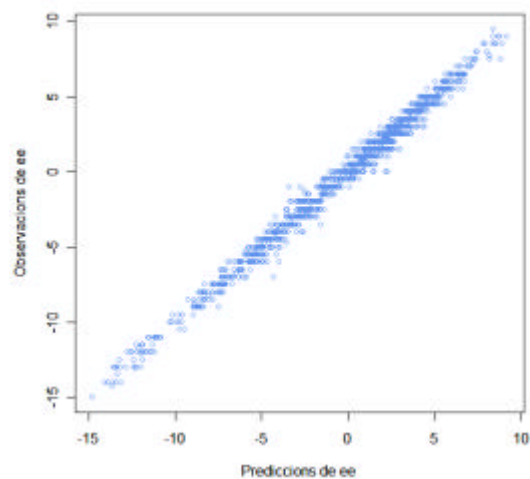


Figura 28: Gràfic dels valors predits vers les observacions del model

En la **figura 27**, observem que ja no tenim pràcticament errors grans. En el gràfic dels valors observats amb les prediccions (**figura 28**) veiem que les hipòtesis de linealitat es compleixen.

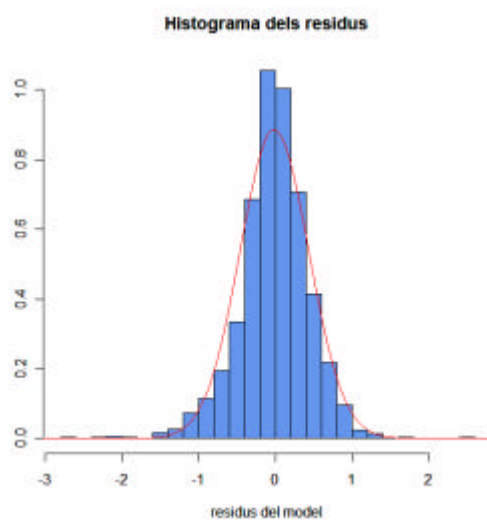


Figura 29: Histograma dels residus del model amb la corba de la normal

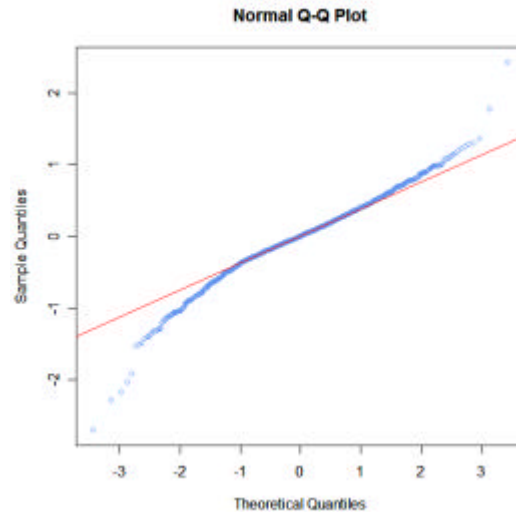


Figura 30: Qqnorm dels residus del model

Les hipòtesis de normalitat han millorat a l'eliminar els valors anòmals. La **figura 30** ja no té les cues tant pesants.

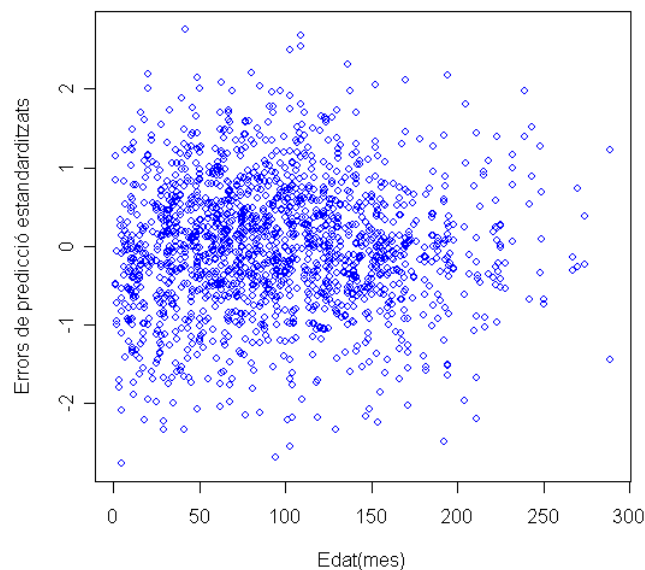


Figura 31: Scatter-plot dels errors de predicció estandarditzats vers l'edat en mesos

En el gràfic anterior, veiem que la variància és més constant que en la **figura 24**, tot i que encara tenim una variabilitat més gran en edats primerenques. Això ens indica que el comportament lineal assumit no és del tot correcte per edats recents. Els experts confirmen que en la primera etapa del creixement la variabilitat observada és molt més gran.

4.6. Models lineals mixtos sense atípics afegint covariables

Model afegint la covariable sexe

Com que la nostra idea és intentar modelitzar amb la màxima presició possible, anem a veure que passa quan introduïm la covariable que fa referència al sexe del pacient. Si la covariable és significativa, podem modelar les dades a partir de dues subpoblacions.

Linear mixed-effects model fit by REML

Random effects:

Formula: ~mes | id

Structure: Diagonal

(Intercept) mes

StdDev: **3.616731 0.01885748**

Formula: ~mes | eye %in% id

Structure: General positive-definite

StdDev Corr

(Intercept) **0.398764335** (Intr)

mes **0.004986847 1**

Residual 1.685315941

Variance function:

Structure: Constant plus power of variance covariate

Formula: ~mes

Parameter estimates:

const power

0.2114085 -0.5656607

Fixed effects: ee ~ mes * sexe

Value Std.Error DF t-value p-value

(Intercept) **1.2961861** 0.3576685 1284 3.623987 0.0003

mes **-0.0029783** 0.0019684 1284 -1.513071 **0.1305**

```

sexel      -0.6170497 0.5013667 214 -1.230735 0.2198
mes:sexel  -0.0001686 0.0027749 1284 -0.060743 0.9516

```

Correlation:

```

(Intr) mes      sexel
mes            -0.035
sexel         -0.713  0.025
mes:sexel      0.025 -0.709 -0.038

```

Standardized Within-Group Residuals:

```

          Min          Q1          Med          Q3          Max
-3.070017133 -0.494811623 -0.007479974  0.498463797  2.711898886

```

Number of Observations: 1718

Number of Groups:

```

id eye %in% id
216      432

```

El model obtingut és:

$$ee_{ijk} = (1.2962 - 0.6170 \cdot I_d + b_{0i} + b_{0ij}) + (-0.0030 - 0.0002 \cdot I_d + b_{1i} + b_{1ij}) mes_{ijk} + e_{ijk}$$

on I_d és un indicador que val 1 en cas que el pacient sigui una dona.

Nivell nen:

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)}) \quad V^{(1)} = \begin{pmatrix} 3.6167 & 0 \\ 0 & 0.0189 \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Nivell ull:

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)}) \quad V^{(2)} = \begin{pmatrix} 0.3988 & 0 \\ 0 & 0.0050 \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Residus:

$$e_{ijk} \sim N(0, (1.6853)^2 (0.2114 + mes^{-0.5657}))$$

Realitzem el test de Likelihood per comprovar si el model millora.

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
	model.lme4.anova	1 10	5011.664	5066.153	-2495.832			
	model.lme5.anova	2 12	5014.139	5079.526	-2495.070	1 vs 2	1.524847	0.4665

Podem veure que el nostre model no millora significativament. Si observem la comparació dels models, veiem que la variable sexe no influeix en la modelització de l'equivalent esfèric del pacient. Per tant, aquesta variable no formarà part del model final.

Model afegint la covariable malaltia a la primera visita

Anem a veure que passa quan introduïm la covariable que fa referència a la malaltia de l'ull en la primera visita.

Linear mixed-effects model fit by REML

Random effects:

Formula: ~mes | id

Structure: Diagonal

(Intercept) mes

StdDev: 1.682055 0.01779148

Formula: ~mes | eye %in% id

Structure: General positive-definite

StdDev Corr

(Intercept) 0.367914759 (Intr)

mes 0.005526334 1

Residual 1.809310947

Variance function:

Structure: Constant plus power of variance covariate

Formula: ~mes

Parameter estimates:

const power

0.2250422 -0.6847714

Fixed effects: ee ~ mes * enf

Value Std.Error DF t-value p-value

(Intercept) -3.226522 0.21476091 1283 -15.023784 0.0000

mes	-0.006437	0.00230819	1283	-2.788982	0.0054
enf0	4.787124	0.23729360	214	20.173846	0.0000
enf1	6.974213	0.28545300	214	24.432087	0.0000
mes:enf0	0.003472	0.00258981	1283	1.340588	0.1803
mes:enf1	0.006437	0.00300744	1283	2.140318	0.0325
Correlation:					
	(Intr)	mes	enf0	enf1	ms:nf0
mes		-0.163			
enf0		-0.766	0.167		
enf1		-0.718	0.132	0.649	
mes:enf0		0.165	-0.772	-0.218	-0.130
mes:enf1		0.137	-0.730	-0.138	-0.162
				0.686	
Standardized Within-Group Residuals:					
	Min	Q1	Med	Q3	Max
	-2.75755952	-0.48260807	-0.01454817	0.49714392	3.16339738
Number of Observations: 1718					
Number of Groups:					
	id	eye	%in%	id	
	216			432	

El model obtingut és:

$$ee_{ijk} = (-3.2265 + 4.7871 \cdot I_e + 6.9742 \cdot I_h + b_{0i} + b_{0ij}) + \\ + (-0.0064 + 0.0035 \cdot I_e + 0.0064 \cdot I_h + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$$

on I_e és un indicador que val 1 en cas que el pacient sigui emmetrop a la primera visita i I_h és un indicador que val 1 en cas que el pacient sigui hipermetrop a la primera visita.

Nivell nen:

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)}C^{(1)}V^{(1)}) \quad V^{(1)} = \begin{pmatrix} 1.6821 & 0 \\ 0 & 0.0178 \end{pmatrix} \quad C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Nivell ull:

$$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)}) \quad V^{(2)} = \begin{pmatrix} 0.3679 & 0 \\ 0 & 0.0055 \end{pmatrix} \quad C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Residus:

$$e_{ijk} \sim N(0, (1.8093)^2 (0.2250 + mes^{-0.6848}))$$

Comprovem si el model ajustat ens millora introduint la covariable malaltia.

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
	model.lme4.anova	1 10	5011.664	5066.153	-2495.832			
	model.lme6.anova	2 14	4696.394	4772.679	-2334.197	1 vs 2	323.2703	<.0001

Podem veure que el nostre model millora significativament. Per tant, la variable malaltia a la primera visita influeix en la modelització de l'evolució posterior del pacient.

A continuació, podem veure la representació de sis ulls (tres nens).

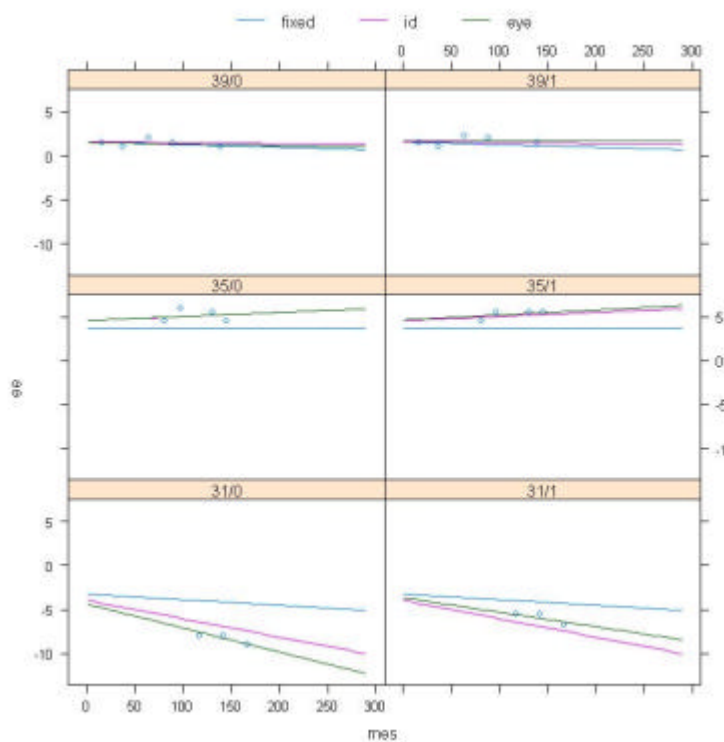


Figura 32: Modelització per 3 individus representatius de la mostra

Els nens que tenim representats en aquest gràfic són els mateixos que en la **figura 25**. Igual que en el gràfic esmentat, la línia de color blau representa la tendència poblacional, en color violeta podem veure la tendència del nen, i en color verd la tendència de l'ull del nen. Aquesta última línia veiem que s'ajusta molt bé amb les observacions d'aquest ull.

A diferència de la **figura 25**, podem veure que la tendència poblacional és diferent per cada un dels individus. Això ve donat per el coeficient que hem introduït al nostre model, i que depèn de la malaltia a la primera visita.

L'individu 31 és representatiu dels miops. Podem veure que hi ha una tendència poblacional pels nens que tenen miopia a la primera visita, aquesta tendència ens posa de manifest que a mesura que avança l'edat, la miopia s'agreuja.

L'individu 35 és emmetrop a la primera visita, però observem que a mesura que avança l'estudi aquest nen ha anat hipermetropitzant cada vegada més. Aquesta no és la tendència de la població emmetrop, ja que aquesta tendeix a mantenir-se estable al llarg del temps.

L'individu 39 és representatiu dels hipermetrops. A l'igual que la població a la que representa, aquest individu ha anat guanyant agudesa visual a mesura que creixia. Podríem dir que aquesta població acostuma a curar-se a mesura que es fa gran.

Podem trobar aquest mateix gràfic per tots els nens d'aquesta mostra a l'**annex 7.4**.

4.7. Resum dels models

MODEL	MILLORES INTRODUIDES	SINTAXI de R	OBSERVACIONS
Model lineal simple		<code>lm(ee~mes, dades)</code>	No és correcte.
Model lineal simple a nivell de nen	Afegim un nivell d'agrupació per la variable id.	<code>lmList(dades)</code>	No és correcte.
Model lineal simple a nivell de nen i ull	Afegim un nivell d'agrupació per la variable eye.	<code>lmList(dades)</code>	Model correcte.
Model lineal mixt 1	Aplicuem model lineal mixt general per fer inferència.	<code>lme(dades, random=list(id=pdSymm(~mes), eye=pdSymm(~mes)))</code>	Model correcte.
Model lineal mixt 2	Canviem a estructura diagonal als efectes aleatoris associats a la variable id.	<code>lme(dades, random=list(id=pdDiag(~mes), eye=pdSymm(~mes)))</code>	Model millorat.
Model lineal mixt 3	Canviem a estructura diagonal als efectes aleatoris associats a la variable eye.	<code>lme(dades, random=list(id=pdDiag(~mes), eye=pdDiag(~mes)))</code>	El model no millora.
Model lineal mixt 4	Aplicuem una potència a la variància del model lineal mixt 2.	<code>lme(dades, random=list(id=pdDiag(~mes), eye=pdSymm(~mes)), weights=varPower(form=~mes), control=list(max.iter=200))</code>	Model millorat.
Model lineal mixt 5	Afegim un nou paràmetre a la variància del model.	<code>lme(dades, random=list(id=pdDiag(~mes), eye=pdSymm(~mes)), weights=varConstPower(form=~mes), control=list(max.iter=200))</code>	Model millorat.
Model lineal mixt sense atípics 1	Eliminem els atípics de les dades.	<code>lme(dades, random=list(id=pdDiag(~mes), eye=pdSymm(~mes)), weights=varConstPower(form=~mes), control=list(max.iter=200))</code>	Model millorat.
Model lineal mixt sense atípics 2	Afegim la covariable sexe a la part fixa del model.	<code>lme(ee~mes*sexe, dades, random=list(id=pdDiag(~mes), eye=pdSymm(~mes)), weights=varConstPower(form=~mes), control=list(max.iter=200))</code>	El model no millora.
Model lineal mixt sense atípics 3	Afegim la covariable enf a la part fixa del model lineal mixt sense atípics 1.	<code>lme(ee~mes*enf, dades, random=list(id=pdDiag(~mes), eye=pdSymm(~mes)), weights=varConstPower(form=~mes), control=list(max.iter=200))</code>	Model millora significativament.

MODEL	EQUACIÓ DEL MODEL	NIVELL NEN	NIVELL ULL	RESIDUS
Model lineal mixt 1	$ee_{ijk} = (0.7830 + b_{0i} + b_{0ij}) + (-0.0025 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 3.6046 & 0 \\ 0 & 0.0208 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0.101 \\ 0.101 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.6631 & 0 \\ 0 & 0.0041 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (0.6073)^2 1)$
Model lineal mixt 2	$ee_{ijk} = (0.7844 + b_{0i} + b_{0ij}) + (-0.0026 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 3.6251 & 0 \\ 0 & 0.0210 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.6642 & 0 \\ 0 & 0.0041 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (0.6062)^2 1)$
Model lineal mixt 3	$ee_{ijk} = (0.7850 + b_{0i} + b_{0ij}) + (-0.0026 + b_{1i} + b_{1ij})...mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 3.6243 & 0 \\ 0 & 0.0209 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.7976 & 0 \\ 0 & 0.0062 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (0.5982)^2 1)$
Model lineal mixt 4	$ee_{ijk} = (0.8272 + b_{0i} + b_{0ij}) + (-0.0032 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 3.7250 & 0 \\ 0 & 0.0178 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.7408 & 0 \\ 0 & 0.0036 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (1.7591)^2 mes^{-0.2536})$
Model lineal mixt 5	$ee_{ijk} = (0.8356 + b_{0i} + b_{0ij}) + (-0.0033 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 3.7274 & 0 \\ 0 & 0.0176 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.7515 & 0 \\ 0 & 0.0035 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 0.996 \\ 0.996 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (4.2328)^2 (0.1043 + mes^{0.8294}))$
Model lineal mixt sense atípics 1	$ee_{ijk} = (0.9823 + b_{0i} + b_{0ij}) + (-0.0031 + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 3.6210 & 0 \\ 0 & 0.0188 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.3999 & 0 \\ 0 & 0.0050 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 0.993 \\ 0.993 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (1.7040)^2 (0.2094 + mes^{0.5693}))$
Model lineal mixt sense atípics 2	$ee_{ijk} = (1.296 - 0.617 \cdot I_d + b_{0i} + b_{0ij}) + (-0.003 - 0.0002 \cdot I_d + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 3.6167 & 0 \\ 0 & 0.0189 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.3988 & 0 \\ 0 & 0.0050 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (1.6853)^2 (0.2114 + mes^{0.5653}))$
Model lineal mixt sense atípics 3	$ee_{ijk} = (-3.227 + 4.787 \cdot I_s + 6.974 \cdot I_h + b_{0i} + b_{0ij}) + (-0.006 + 0.003 \cdot I_s + 0.006 \cdot I_h + b_{1i} + b_{1ij})mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 1.6821 & 0 \\ 0 & 0.0178 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0ij} \\ b_{1ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.3679 & 0 \\ 0 & 0.0055 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (1.8093)^2 (0.2250 + mes^{0.6844}))$

4.8. Interpretacions

Al llarg de l'anàlisi de les dades, hem pogut veure la gran variabilitat que hi ha entre els nens de la mostra. Per reduir aquesta variabilitat, i a la vegada aconseguir un model més robust, hem eliminat els individus amb una i dues visites.

En el primer model vàlid que hem trobat (model lineal amb dades agrupades a dos nivell), la variància residual és de $(0.6033)^2$. Cal recordar que aquest és un model lineal general saturat, per tant, no aconseguirem millorar aquest valor, excepte amb la supressió dels atípics.

Al primer model lineal mixt que hem ajustat, la variància per la part aleatòria a nivell nen és de $(3.6046)^2$ per l'intercep i de $(0.0208)^2$. Tot i que, aquest model és més senzill que el final, no ajusta tant bé les dades, i no ens permet fer un diagnòstic de cara a l'evolució de la malaltia tan acurat com en el model final. En el model final abans de treure les dades atípiques, la variància de la part aleatòria del model a nivell nen és de $(3.7274)^2$ per l'intercep i de $(0.0176)^2$.

Hem pogut definir un model amb menys variància, perquè hem tret els valors atípics. Aquest model, ens queda amb una variància de $(3.6210)^2$ per l'intercep i de $(0.0188)^2$. Hem ajustat aquest model final sense atípics afegint-hi la covariable sexe, i hem vist que el model no millorava. En canvi, i com ja s'intuïa, la malaltia que pateix el nen a la primera visita és important a la hora de determinar el seu pronòstic evolutiu. En aquest model amb la covariable malaltia, la variància de la part aleatòria del model a nivell nen és de $(1.6821)^2$ per l'intercep i de $(0.0178)^2$.

A les nostres dades, els valors atípics poden venir donats per 3 casos:

- errors en alguns dels mesurament de l'equivalent esfèric,
- comportament anòmal en l'evolució d'algun ull del nen, ja sigui per una operació o per un agreujament molt ràpid de la malaltia,
- i comportament diferent en els dos ulls.

Es pot concloure, que la població de la mostra és una mixtura en funció de la malaltia a la primera visita. Tenim tres subgrups en funció de la malaltia a la primera visita i cada una d'elles té una evolució diferent. En canvi, la variable sexe no determina cap subpoblació. A més a més, el principal avantatge d'aquest model és que permet fer prediccions sobre l'evolució del pacient en funció de l'edat i la malaltia a l'inici.

5. Conclusions

En aquest projecte hem volgut aprofundir en l'estudi de les dades amb mesures repetides al llarg del temps.

Existeixen diferents models per analitzar les mesures repetides al llarg del temps.

Durant la carrera, hem vist els models **ANOVA per a mesures repetides** i els models **MANOVA**. Un dels principals inconvenients que trobem quan apliquem aquests models, és que les dades han d'estar balancejades i mesurades en el mateix instant de temps.

Quan les dades no segueixen el patró descrit anteriorment, no podem aplicar els models lineals clàssics. Aquests models requereixen que les dades segueixin un mostra aleatòria simple, i normalment les dades amb mesures repetides estan correlacionades a diferents nivells. Si volguéssim ajustar un disseny per aquest model, podríem ajustar un model per a cada individu però aconseguiríem un model sobreparametritzat. Si ajustéssim un model poblacional, no tindríem en compte la correlació entre les observacions d'un mateix individu.

La metodologia que hem de seguir per analitzar el conjunt de dades, com el de l'exemple, és la dels models lineal mixtos. Els models lineals mixtos ens permeten modelitzar dades no balancejades i mesurades en diferents instants de temps. Existeixen diferents maneres de formular els models lineals mixtos, cosa que ens permet treballar amb les dades de forma diferent: una d'elles és la que ens permet que la part aleatòria inclogui informació de cada individu. En un altre model, cada individu comparteix una part de la informació poblacional a la vegada que té un coeficient aleatori per a cada individu. Un tercer model, que s'acostuma a utilitzar en mesures biomèdiques i sociològiques, modelitza a dos nivells, els coeficients i les observacions.

Per fer més senzilla la comprensió de la part teòrica, hem volgut il·lustrar amb un cas pràctic els models lineals mixtos. El conjunt de dades utilitzat és més complex perquè té les dades agrupades a dos nivells. Per poder millorar l'ajust del model, hem treballat l'estructura en la matriu de variàncies i covariàncies dels efectes aleatoris,

hem modelitzat la heteroscedasticitat dels residus i l'estructura en la matriu de correlacions. En el conjunt de dades aplicat, hem pogut treballar les dades atípiques, tant a nivell d'observació com d'unitat experimental.

L'objectiu d'aquest projecte ha estat aprofundir en el coneixement dels models lineals mixtos a la vegada que s'il·lustrava amb un exemple a mode de tutorial. D'aquesta manera, resulta més senzilla l'aplicació d'aquesta metodologia en altres àmbits.

5.1. Possibles extensions

Existeixen extensions d'aquests models per a casos on la relació entre les dades no sigui lineal. La relació d'aquestes dades podria tenir tendència asimptòtica, quadràtica, polinòmica... En aquests casos podem aplicar els models no lineals mixtos.

Ens podem trobar cassos on la variable resposta segueixi una distribució binomial, una distribució poisson... En resum, en el cas que la variable resposta no segueixi una distribució gaussiana podem aplicar els models lineals generalitzats mixtos.

Per un conjunt de dades similar al analitzat, una altra extensió dels models lineals mixtos, la podríem trobar quan introduïm variables que depenen del temps. Aquesta extensió no l'hem pogut aplicar perquè, la variable resposta del conjunt de dades que hem utilitzat per aplicar els models lineals mixtos és lineal, segueix una distribució gaussiana i, les variables sexe i malaltia són fixes (no depenen dels temps).

6. Bibliografia

Llibres:

- W.E. Nelson, V.C. Vaughan III i K.J.McKay (Vuitena edició) *Tratado de pediatria*, Salvat (Volum I i Volum II).
- M.A. Fernández, B. Mingo, R. Rodríguez, M^a.D. Torres, *Gaia-3 (ciencias naturales)*, Vicens Vives.
- ¹ Pinheiro, J.C., Bates, D.M. (2000) *Mixed-effect Models in S and S-Plus*, Springer-Verlag.
- Geert Verbeke, Geert Molenberghs, *Linear Mixed Models in Practice*, Springer.
- W. N. Venables and B. D. Ripley, (2002) *Modern Applied Statistic with S*, Springer.

Pàgines web:

Relacionades amb el coneixement general de la síndrome de down

Fundació Síndrome de Down de Cantabria

<http://www.infonegocio.com/downcan/todo/curso/actualizacion/cuidsalud.htm>

La Asociación Síndrome de Down de Cádiz y Bahía "Lejeune"

<http://www2.uca.es/huesped/down/introdown.htm>

Fundació catalana per la SD

<http://www.fcscd.org/cas/revista/downloads.htm>

Congrés internacional de SD

<http://www.ndsccenter.org/>

http://www.down21.org/vision_perspec/art_que_es_sd.htm

http://es.wikipedia.org/wiki/S%C3%ADndrome_de_Down

Relacionades amb l'oftalmologia

<http://uuhsc.utah.edu/healthinfo/spanish/Eye/glossary.htm>

<http://es.wikipedia.org/wiki/Ojo>

<http://www.geocities.com/cirugialaser/InfoDefectos.htm#miopia>

<http://www.horusgo.com/elojo.htm>

http://www.tuotromedico.com/temas/alteraciones_ojo_miopia.htm

http://www.cverges.com/castellano/INFORMACION/enfermedades/errores_refractivos.htm

<http://www.cverges.com/castellano/INFORMACION/enfermedades/nistagmus.htm>

<http://www.misojosonline.com.ve/corr.htm> <http://www.colegiodeopticos.cl/e.html>

Relacionada amb el síndrome down i la visió

<http://www.imagenoptica.com.mx/pdf/revista28/1.pdf>

Diccionaris de termes mèdics

<http://www.iqb.es/diccio/b/bl.htm>

<http://www.medic.cat/>

<http://dlc.iec.cat/>

www.wikipedia.com

<http://isi.cbs.nl/glossary/index.htm>

<http://www.nlm.nih.gov/medlineplus/ency/article/000997.htm>

Articles d'oftalmologia on mesuren dades semblants a les nostres

Relación entre los componentes ópticos oculares e implicaciones en el proceso de emetropización

<http://www.oftalmo.com/seo/archivos/articulo.php?idSolicitud=2271&numR=5&mesR=5&anioR=2008&idR=135>

Miopía inducida por la pilocarpina

<http://www.oftalmo.com/seo/archivos/articulo.php?idSolicitud=297&numR=2&mesR=2&anioR=1998&idR=2>

Patología oftalmológica comparada entre dos poblaciones no seleccionadas: ayacucho-perú y sama de langreo-españa

<http://www.oftalmo.com/seo/archivos/articulo.php?idSolicitud=298&numR=2&mesR=2&anioR=1998&idR=2>

³ Development of refractive error and strabismus in children with down syndrome

<http://www.iovs.org/cgi/reprint/44/3/1023>

Olav H Haugen, Gunnar Høvdning, Isa Lundström

² Refractive development in children with down's syndrome: a population based, longitudinal study

<http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1723994&blobtype=pdf>

Mary Cregg, J. Margaret Woodhouse i altres

Articles relacionats amb la Màxima versemblança

<http://www.seh-lelha.org/maxverosim.htm>

http://statgen.iop.kcl.ac.uk/bqim/mle/sslike_1.html

Análisis de estudios longitudinales, datos agrupados y medidas repetidas

<http://www.seh-lelha.org/mixedmodels.htm>

Altres pàgines web:

<http://www2.uca.es/serv/ai/formacion/spss/Pantalla/16anovar.pdf>

<http://grad.uprm.edu/tesis/gaonabarrera.pdf>

<http://eio.usc.es/eipc1/MATERIALES/091561963.pdf>

<http://books.google.es/books?id=xjwsfuRXIjIC&pg=PA197&lpg=PA197&dq=test+de+wald&source=web&ots=IHxEMXrUTg&sig=jFN2PSCK0qCDCWUOBxXy4uXNgcY&hl=es#PPA174,M1>

7. Annex

7.1. Sintaxi utilitzada al programa SPSS (versió 15.0)

```
*****
* FITXER DownTot2tot.txt *.
*****

* Recuperem el fitxer amb les dades original.
GET DATA /TYPE = TXT
/FILE = 'F:\projecte\DownTot2tot.txt'
/DELCASE = LINE
/DELIMITERS = "\t"
/ARRANGEMENT = DELIMITED
/FIRSTCASE = 2
/IMPORTCASE = ALL
/VARIABLES =
id F1.0
visit F1.0
Nvisits F1.0
mes F3.0
ini F3.0
period F3.0
eye F1.0
ee F6.3
ast F4.2
sexe F1.0
enf F2.0
evol F2.0
.
CACHE.
EXECUTE.
DATASET NAME DownTot2tot WINDOW=FRONT.
EXECUTE.

VARIABLE LABEL SEXE 'Sexe del pacient'.
VALUE LABELS SEXE 0 'Home' 1 'Dona'.
EXECUTE.

VARIABLE LABEL ENF 'Malaltia del pacient a la primera visita'.
```

```
VALUE LABELS ENF -1 'Miopia' 0 'Emmetropia' 1 'Hipermetropia'.  
EXECUTE.
```

```
VARIABLE LABEL EVOL 'Evolució del pacient'.  
VALUE LABELS EVOL -1 'Decreixent' 0 'Estable' 1 'Creixent'.  
EXECUTE.
```

```
VARIABLE LABEL NVisits 'Nombre de visites'.  
VARIABLE LABEL Ini 'Edat a la primera visita'.  
VARIABLE LABEL Period 'Temps de seguiment'.  
VARIABLE LABEL Ee 'Equivalent esfèric'.  
VARIABLE LABEL Ast 'Astigmatisme'.  
EXECUTE.
```

```
* Identificar casos duplicados.  
SORT CASES BY id(A) .  
MATCH FILES /FILE = * /BY id  
/FIRST = PrimarioPrimero /LAST = PrimarioÚltimo.  
DO IF (PrimarioPrimero).  
COMPUTE SecuenciaCoincidencia = 1 - PrimarioÚltimo.  
ELSE.  
COMPUTE SecuenciaCoincidencia = SecuenciaCoincidencia + 1.  
END IF.  
LEAVE SecuenciaCoincidencia.  
FORMAT SecuenciaCoincidencia (f7).  
COMPUTE InDupGrp = SecuenciaCoincidencia > 0.  
SORT CASES InDupGrp(D).  
MATCH FILES /FILE = * /DROP = PrimarioÚltimo InDupGrp SecuenciaCoincidencia.  
VARIABLE LABELS PrimarioPrimero 'Indicador de cada primer caso coincidente'+  
' como Primario' .  
VALUE LABELS PrimarioPrimero 0 'Caso duplicado' 1 'Caso primario'.  
VARIABLE LEVEL PrimarioPrimero (ORDINAL).  
FREQUENCIES VARIABLES = PrimarioPrimero .  
EXECUTE.
```

```
* Filtrem primer cas de cada nen.  
USE ALL.  
COMPUTE filter_$=(PrimarioPrimero = 1).  
VARIABLE LABEL filter_$ 'PrimarioPrimero = 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'No seleccionado' 1 'Seleccionado'.
```

```

FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .

```

* Traiem les taules de freqüències de les variables.

```

FREQUENCIES
  VARIABLES=sexe enf
  /PIECHART PERCENT
  /ORDER= ANALYSIS .
EXECUTE.

```

```

FREQUENCIES
  VARIABLES=NVisits
  /BARCHART FREQ
  /ORDER= ANALYSIS .
EXECUTE.

```

```

EXAMINE
  VARIABLES=ini period
  /PLOT NONE
  /PERCENTILES(25,50,75) HAVERAGE
  /STATISTICS DESCRIPTIVES
  /MISSING LISTWISE
  /NOTOTAL.
EXECUTE.

```

* Traiem el filtre.

```

FILTER OFF.
USE ALL.
EXECUTE .

```

* Identificar casos duplicados.

```

SORT CASES BY id(A) eye(A) .
MATCH FILES /FILE = * /BY id eye
  /FIRST = PrimarioPrimero1 /LAST = PrimarioÚltimo.
DO IF (PrimarioPrimero1).
COMPUTE SecuenciaCoincidencia = 1 - PrimarioÚltimo.
ELSE.
COMPUTE SecuenciaCoincidencia = SecuenciaCoincidencia + 1.
END IF.

```

```
LEAVE SecuenciaCoincendencia.  
FORMAT SecuenciaCoincendencia (f7).  
COMPUTE InDupGrp = SecuenciaCoincendencia > 0.  
SORT CASES InDupGrp(D).  
MATCH FILES /FILE = * /DROP = PrimarioÚltimo InDupGrp SecuenciaCoincendencia.  
VARIABLE LABELS PrimarioPrimero1 'Indicador de cada primer caso coincidente' +  
  ' como Primario' .  
VALUE LABELS PrimarioPrimero1 0 'Caso duplicado' 1 'Caso primario'.  
VARIABLE LEVEL PrimarioPrimero1 (ORDINAL).  
FREQUENCIES VARIABLES = PrimarioPrimero1 .  
EXECUTE.
```

* Filtrem primer cas de cada nen.

```
USE ALL.  
COMPUTE filter_$=(PrimarioPrimero = 1).  
VARIABLE LABEL filter_$ 'PrimarioPrimero = 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'No seleccionado' 1 'Seleccionado'.  
FORMAT filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE .
```

* Traiem les taules de freqüències de les variables.

```
FREQUENCIES  
  VARIABLES=evol  
  /PIECHART PERCENT  
  /ORDER= ANALYSIS .  
EXECUTE.
```

```
FREQUENCIES  
  VARIABLES=ast  
  /BARCHART FREQ  
  /ORDER= ANALYSIS .  
EXECUTE.
```

```
EXAMINE  
  VARIABLES=ee ast  
  /PLOT NONE  
  /PERCENTILES(25,50,75) HAVERAGE  
  /STATISTICS DESCRIPTIVES  
  /MISSING LISTWISE
```


/NOTOTAL.

EXECUTE.

* FITXER DownTot23visites.txt **.

* Recuperem el fitxer amb les dades original.

GET DATA /TYPE = TXT

/FILE = 'F:\projecte\DownTot23visites.txt'

/DELCASE = LINE

/DELIMITERS = "\t"

/ARRANGEMENT = DELIMITED

/FIRSTCASE = 2

/IMPORTCASE = ALL

/VARIABLES =

id F1.0

visit F1.0

Nvisits F1.0

mes F3.0

ini F3.0

period F3.0

eye F1.0

ee F6.3

ast F4.2

sexe F1.0

enf F2.0

evol F2.0

.

CACHE.

EXECUTE.

DATASET NAME DownTot23visites WINDOW=FRONT.

EXECUTE.

VARIABLE LABEL SEXE 'Sexe del pacient'.

VALUE LABELS SEXE 0 'Home' 1 'Dona'.

EXECUTE.

VARIABLE LABEL ENF 'Malaltia del pacient a la primera visita'.

VALUE LABELS ENF -1 'Miopia' 0 'Emmetropia' 1 'Hipermetropia'.

EXECUTE.

```
VARIABLE LABEL EVOL 'Evolució del pacient'.
VALUE LABELS EVOL -1 'Decreixent' 0 'Estable' 1 'Creixent'.
EXECUTE.
```

```
VARIABLE LABEL NVisits 'Nombre de visites'.
VARIABLE LABEL Ini 'Edat a la primera visita'.
VARIABLE LABEL Period 'Temps de seguiment'.
VARIABLE LABEL Ee 'Equivalent esfèric'.
VARIABLE LABEL Ast 'Astigmatisme'.
EXECUTE.
```

```
* Identificar casos duplicados.
SORT CASES BY id(A) .
MATCH FILES /FILE = * /BY id
  /FIRST = PrimarioPrimero /LAST = PrimarioÚltimo.
DO IF (PrimarioPrimero).
  COMPUTE SecuenciaCoincidencia = 1 - PrimarioÚltimo.
ELSE.
  COMPUTE SecuenciaCoincidencia = SecuenciaCoincidencia + 1.
END IF.
LEAVE SecuenciaCoincidencia.
FORMAT SecuenciaCoincidencia (f7).
COMPUTE InDupGrp = SecuenciaCoincidencia > 0.
SORT CASES InDupGrp(D).
MATCH FILES /FILE = * /DROP = PrimarioÚltimo InDupGrp SecuenciaCoincidencia.
VARIABLE LABELS PrimarioPrimero 'Indicador de cada primer caso coincidente'+
  ' como Primario' .
VALUE LABELS PrimarioPrimero 0 'Caso duplicado' 1 'Caso primario'.
VARIABLE LEVEL PrimarioPrimero (ORDINAL).
FREQUENCIES VARIABLES = PrimarioPrimero .
EXECUTE.
```

```
* Filtrem primer cas de cada nen.
USE ALL.
COMPUTE filter_$=(PrimarioPrimero = 1).
VARIABLE LABEL filter_$ 'PrimarioPrimero = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'No seleccionado' 1 'Seleccionado'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .
```

* Traiem les taules de freqüències de les variables.

FREQUENCIES

VARIABLES=sexe enf

/PIECHART PERCENT

/ORDER= ANALYSIS .

EXECUTE.

FREQUENCIES

VARIABLES=NVisits

/BARChart FREQ

/ORDER= ANALYSIS .

EXECUTE.

EXAMINE

VARIABLES=ini period

/PLOT NONE

/PERCENTILES(25,50,75) HAVERAGE

/STATISTICS DESCRIPTIVES

/MISSING LISTWISE

/NOTOTAL.

EXECUTE.

* Traiem el filtre.

FILTER OFF.

USE ALL.

EXECUTE .

* Identificar casos duplicados.

SORT CASES BY id(A) eye(A) .

MATCH FILES /FILE = * /BY id eye

/FIRST = PrimarioPrimero1 /LAST = PrimarioÚltimo.

DO IF (PrimarioPrimero1).

COMPUTE SecuenciaCoincidencia = 1 - PrimarioÚltimo.

ELSE.

COMPUTE SecuenciaCoincidencia = SecuenciaCoincidencia + 1.

END IF.

LEAVE SecuenciaCoincidencia.

FORMAT SecuenciaCoincidencia (f7).

COMPUTE InDupGrp = SecuenciaCoincidencia > 0.

SORT CASES InDupGrp(D).

```
MATCH FILES /FILE = * /DROP = PrimarioÚltimo InDupGrp SecuenciaCoincidencia.
VARIABLE LABELS PrimarioPrimero1 'Indicador de cada primer caso coincidente' +
' como Primario' .
VALUE LABELS PrimarioPrimero1 0 'Caso duplicado' 1 'Caso primario'.
VARIABLE LEVEL PrimarioPrimero1 (ORDINAL).
FREQUENCIES VARIABLES = PrimarioPrimero1 .
EXECUTE.
```

* Filtrem primer cas de cada nen.

```
USE ALL.
COMPUTE filter_$=(PrimarioPrimero = 1).
VARIABLE LABEL filter_$ 'PrimarioPrimero = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'No seleccionado' 1 'Seleccionado'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .
```

* Traiem les taules de freqüències de les variables.

```
FREQUENCIES
  VARIABLES=evol
  /PIECHART PERCENT
  /ORDER= ANALYSIS .
EXECUTE.
```

```
FREQUENCIES
  VARIABLES=ast
  /BARChart FREQ
  /ORDER= ANALYSIS .
EXECUTE.
```

```
EXAMINE
  VARIABLES=ee ast
  /PLOT NONE
  /PERCENTILES(25,50,75) HAVERAGE
  /STATISTICS DESCRIPTIVES
  /MISSING LISTWISE
  /NOTOTAL.
EXECUTE.
```

7.2. Sintaxi utilitzada al programa R (versió 2.3.0)

```
# Llibreria per construir Models Lineals Mixtos amb diferents nivells
library(nlme)
# Les dades estan a ChildDown2.exe
# arxiu amb totes les visites (3064 registres i 689 individus)
# dades<-read.table("F:\\projecte\\DownTot2tot.txt", header=T,dec=",")

# arxiu amb 2 visites o mes (2442 registres i 378 individus)
# dades<-read.table("F:\\projecte\\DownTot22visites.txt", header=T,dec=",")

# arxiu amb 3 visites o mes (1846 registres i 229 individus)
dades<-read.table("F:\\projecte\\DownTot23visites.txt", header=T,dec=",")

# arxiu amb 3 visites o mes i sense atípics (1718 registres i 216 individus)
# dades<-read.table("F:\\projecte\\DownTot23visites-noatipics.txt", header=T,dec=",")
)

id <-dades$id
visit <-dades$visit
Nvisits <-dades$Nvisits
mes <-dades$mes
ini <-dades$ini
period <-dades$period
eye <-dades$eye
ee <-dades$ee
ast <-dades$ast
sexe <-as.factor(dades$sexe)
enf <-as.factor(dades$enf)
evol <-as.factor(dades$evol)
dades$sexe <-as.factor(dades$sexe)
dades$enf <-as.factor(dades$enf)
dades$evol <-as.factor(dades$evol)

# Els descriptius al final els hem fet amb el programa SPSS.
# summary(dades)
# win.graph()
# par(mfrow=c(1,3))
# hist(Nvisits, col="Cornflower blue", main="Histograma del número de visites",
xlab="número de visites", ylab="n")
# hist(ee, col="Cornflower blue", main="Histograma de l'equivalent esfèric",
xlab="equivalent esfèric", ylab="n")
# hist(ast, col="Cornflower blue", main="Histograma de l'astigmatisme",
xlab="astigmatisme", ylab="n")
# win.graph()
# par(mfrow=c(1,2))
# plot(enf, col="Cornflower blue", main="Histograma de la malaltia de l'ull",
xlab="malaltia de l'ull", ylab="n")
# plot(evol, col="Cornflower blue", main="Histograma de l'evolució de la malaltia de
l'ull", xlab="evolució", ylab="n")
```

```
#####.
# Fem els models que ens interessin.
#####.

# Aquest primer model no és correcte.
model.lm0<-lm(ee~mes, dades)
model.lm0
summary(coef(model.lm0))
anova(model.lm0)

# Aquest model és incorrecte perquè no hi ha independència entre les dades del model
perquè introduïm dades repetides per culpa dels ulls.
dades1=groupedData(ee~mes|id, dades, outer=~sexe)
model.lmList1<-lmList(dades1)
model.lmList1

# Ara les dades ja són independents.
# El model ja és vàlid però té un excés de variables.
# És el model saturat.
dades2=groupedData(ee~mes|id/eye, dades, outer=~sexe)
model.lmList2<-lmList(dades2)
model.lmList2

rane<-ranef(model.lmList2,augFrame=T,level=1)
rane
#plot(rane[1:50,1],rane[1:50,2],main="Gràfic dels ulls dels nens",xlab="Intercept",ylab="pendent")
jpeg(file="F:\\projecte\\grafic ulls nens(model.lmList2).jpg",width=768,height=768)
plot(rane[,1],rane[,2],main="Gràfic dels ulls dels nens",xlab="Intercept",ylab="pendent",col="Cornflower blue",ylim=c(-0.1,0.1))
abline(h=0,col="black",lty=2)
abline(v=0,col="black",lty=2)
abline(v=-3,col="red",lty=3)
abline(v=1,col="red",lty=3)
dev.off()

# Per aconseguir una resolució més gran traiem el gràfic en un fitxer.
jpeg(file="F:\\projecte\\intervalles(model.lmList2).jpg",width=2048,height=2048)
plot(intervals(model.lmList2))
dev.off()

# Primer model mixt amb agrupació a 2 nivells
model.lme31<-lme(dades2,random=list(id=pdSymm(~mes), eye=pdSymm(~mes)))
summary(model.lme31)
intervals(model.lme31)

# Com que cor((Intercept),mes) del id conté el 0, fem pdDiag per simplificar el model
aplicant una estructura a la matriu.
# com que al model anterior hem vist que es pot eliminar una variable (id(~mes)),
fem estructura pdDiag.
model.lme32<-lme(dades2,random=list(id=pdDiag(~mes), eye=pdSymm(~mes)))
```

```

summary(model.lme32)
intervals(model.lme32)

# Igual que en el cas anterior, cor((Intercept),mes) del eye conté el 0, fem pdDiag per
simplificar el model aplicant una estructura a la matriu.
#Mirem a veure si encara podem reduir una mica més els paràmetres.
model.lme33<-lme(dades2,random=list(id=pdDiag(~mes), eye=pdDiag(~mes)))
summary(model.lme33)
intervals(model.lme33)

anova (model.lme32, model.lme33)
# Pel test de likelihood ratio veiem que ens hem de quedar amb el model model-
lme32.
# Perquè el L-ratio és més petit.
# Com que ja tenim un model correcte, anem a fer un estudi dels seus residus per
validar-lo.

# Farem el plot del model indicant els valors extrems.
win.graph()
plot(model.lme32, id=0.01)

win.graph()
model.lme32.pred<-predict(model.lme32)
plot(model.lme32.pred,ee,col="Cornflower blue",ylab="Observacions de
ee",xlab="Prediccions de ee")

win.graph()
hist(resid(model.lme32),breaks=34,col="Cornflower blue",prob=T,xlab="residus del
model",ylab="",main="Histograma dels residus")
curve(dnorm(x,mean=mean(resid(model.lme32)),sd=sd(resid(model.lme32))),col=2,add=T)

win.graph()
qqnorm(resid(model.lme32),col="Cornflower blue")
qqline(resid(model.lme32),col="red")
# Podem veure que tenim cues pesants.

# Fem un gràfic per veure si tenim heteroscedasticitat.
win.graph()
plot(resid(model.lme32,type='n')~mes,xlab="Edat(mes)",ylab="Errors de predicció
estandarditzats",col="blue")
# Podem veure que la variància disminueix a mesura que augmenta l'edat.

# Afegim paràmetres al model per tal d'aconseguir homoscedasticitat.
model.lme34<-lme(dades2,random=list(id=pdDiag(~mes),
eye=pdSymm(~mes)),weights=varPower(form=~mes),control=list(max.iter=200))
model.lme34

# Fem el gràfic de la variància en funció dels mesos.
var.lme34<-function(x){(1.7191^2)+(x^(-0.2536))}

```

```
curve(var.lme34,xlim=c(1,170),col="blue",xlab="mesos",ylab="Variància en funció dels mesos",main="Gràfic de la variància del model 5")
```

```
intervals(model.lme34)
# No podem fer el test de Wald perquè la matriu no és semidefinida positiva, i per tant, no és invertible.
anova(model.lme32,model.lme34)
# Ens quedem amb el model model.lme34.
```

```
# Afegim un altre paràmetre al model.
model.lme35<-lme(dades2,random=list(id=pdDiag(~mes),
eye=pdSymm(~mes)),weights=varConstPower(form=~mes),control=list(max.iter=200))
model.lme35
```

```
# Fem el gràfic de la variància en funció dels mesos.
var.lme35<-function(x){(4.2328^2)*(0.1043+x^(-0.8294))}
curve(var.lme35,xlim=c(1,170),col="blue",xlab="mesos",ylab="Variància en funció dels mesos",main="Gràfic de la variància del model 5")
```

```
intervals(model.lme35)
anova(model.lme34,model.lme35)
```

```
# Farem el plot del model indicant els valors extrems.
win.graph()
plot(model.lme35, id=0.01)
```

```
win.graph()
model.lme35.pred<-predict(model.lme35)
plot(model.lme35.pred,ee,col="Cornflower blue",ylab="Observacions de ee",xlab="Prediccions de ee")
```

```
win.graph()
hist(resid(model.lme35),breaks=34,col="Cornflower blue",prob=T,xlab="residus del model",ylab="",main="Histograma dels residus")
curve(dnorm(x,mean=mean(resid(model.lme35)),sd=sd(resid(model.lme35))),col=2,add=T)
```

```
win.graph()
qqnorm(resid(model.lme35),col="Cornflower blue")
qqline(resid(model.lme35),col="red")
# Podem veure que tenim cues pesants.
```

```
# Fem un gràfic per veure si tenim heteroscedasticitat.
win.graph()
plot(resid(model.lme35,type='n')~mes,xlab="Edat(mes)",ylab="Errors de predicció estandarditzats",col="blue")
# Podem veure que la variància disminueix a mesura que augmenta l'edat.
```



```

# Fem un gràfic on podem veure l'ajust dels diferents ulls amb el model.
plot(augPred(model.lme35,level=0:2),layout=c(6,4,1))

jpeg(file="F:\\projecte\\ajust tots nens(model.lme35).jpg",width=1024,height=6144)
plot(augPred(model.lme35,level=0:2),layout=c(7,66,1))
dev.off()

jpeg(file="F:\\projecte\\ajust nens(model.lme35).jpg",width=2048,height=2048)
plot(augPred(model.lme35,level=0:2),layout=c(6,5,1))
dev.off()

# Fem un gràfic per a tres individus representatius de la mostra.
aa=augPred(model.lme35,level=0:2)
sele=c("39/0","39/1","35/0","35/1","31/0","31/1")
jpeg(file="F:\\projecte\\nens
representatius(model.lme35).jpg",width=768,height=768)
plot(aa[aa[,2]%in%sele,],layout=c(2,3,1))
dev.off()

# Fem un gràfic per a tres individus atípics de la mostra.
sele=c("60/0","60/1","74/0","74/1","85/0","85/1","114/0","114/1","350/0","350/1","3
72/0","372/1","388/0","388/1","578/0","578/1","599/0","599/1","633/0","633/1","672
/0","672/1","703/0","703/1","730/0","730/1")
jpeg(file="F:\\projecte\\nens atipics(model.lme35).jpg",width=768,height=768)
plot(aa[aa[,2]%in%sele,],layout=c(4,7,1))
dev.off()

# Carreguem l'arxiu amb les dades sense els atípics i tornem a modelitzar.
dades.noatip<-read.table("F:\\projecte\\DownTot23visites-noatipics.txt",
header=T,dec=",")
library(lattice)

id <-dades.noatip$id
visit <-dades.noatip$visit
Nvisits <-dades.noatip$Nvisits
mes <-dades.noatip$mes
ini <-dades.noatip$ini
period <-dades.noatip$period
eye <-dades.noatip$eye
ee <-dades.noatip$ee
ast <-dades.noatip$ast
sexe <-as.factor(dades.noatip$sexe)
enf <-as.factor(dades.noatip$enf)
evol <-as.factor(dades.noatip$evol)
dades.noatip$sexe <-as.factor(dades.noatip$sexe)
dades.noatip$enf <-as.factor(dades.noatip$enf)
dades.noatip$evol <-as.factor(dades.noatip$evol)

# Modelitzem amb les dades sense atípics, el model amb millor ajust dels anteriors.
dades3=groupedData(ee~mes|id/eye, dades.noatip, outer=~sexe)

```

```
model.lme4<-lme(dades3,random=list(id=pdDiag(~mes),
eye=pdSymm(~mes)),weights=varConstPower(form=~mes),control=list(max.iter=20
0))
model.lme4
intervals(model.lme4)
anova(model.lme35,model.lme4)

# Farem el plot del model indicant els valors extrems.
win.graph()
plot(model.lme4, id=0.01)

win.graph()
model.lme4.pred<-predict(model.lme4)
plot(model.lme4.pred,ee,col="Cornflower blue",ylab="Observacions de
ee",xlab="Prediccions de ee")

win.graph()
hist(resid(model.lme4),breaks=34,col="Cornflower blue",prob=T,xlab="residus del
model",ylab="",main="Histograma dels residus")
curve(dnorm(x,mean=mean(resid(model.lme4)),sd=sd(resid(model.lme4))),col=2,add
=T)

win.graph()
qqnorm(resid(model.lme4),col="Cornflower blue")
qqline(resid(model.lme4),col="red")
# Podem veure que tenim cues pesants.

# Fem un gràfic per veure si tenim heteroscedasticitat.
win.graph()
plot(resid(model.lme4,type='n')~mes,xlab="Edat(mes)",ylab="Errors de predicció
estandarditzats",col="blue")
# Podem veure que la variància disminueix a mesura que augmenta l'edat.

# Anem a estudiar el model sense atípics i afegint-hi la covariable de sexe.
model.lme5<-lme(ee~mes*sexe,dades3,random=list(id=pdDiag(~mes),
eye=pdSymm(~mes)),weights=varConstPower(form=~mes),control=list(max.iter=20
0))
model.lme5
summary(model.lme5)
intervals(model.lme5)
# No podem comparar els dos models, si hem buscat els estimadors dels paràmetres
per REML.
# Per fer-ho haurem de reformular els models mitjançant Maximum Likelihood.
model.lme5.anova<-lme(ee~mes*sexe,dades3,random=list(id=pdDiag(~mes),
eye=pdSymm(~mes)),weights=varConstPower(form=~mes),control=list(max.iter=20
0),method='ML')
model.lme4.anova<-lme(dades3,random=list(id=pdDiag(~mes),
eye=pdSymm(~mes)),weights=varConstPower(form=~mes),control=list(max.iter=20
0),method='ML')
anova(model.lme4.anova,model.lme5.anova)
```

```

# Anem a estudiar el model sense atípics i afegint-hi la covariable de enf.
model.lme6<-lme(ee~mes*enf,dades3,random=list(id=pdDiag(~mes),
eye=pdSymm(~mes)),weights=varConstPower(form=~mes),control=list(max.iter=20
0))
model.lme6
summary(model.lme6)
intervals(model.lme6)
# No podem comparar els dos models, si hem buscat els estimadors dels paràmetres
per REML.
# Per fer-ho haurem de reformular els models mitjançant Maximum Likelihood.
model.lme6.anova<-lme(ee~mes*enf,dades3,random=list(id=pdDiag(~mes),
eye=pdSymm(~mes)),weights=varConstPower(form=~mes),control=list(max.iter=20
0),method='ML')
anova(model.lme4.anova,model.lme6.anova)

# Fem el gràfic dels nens representatius de la mostra.
aa2=augPred(model.lme6,level=0:2)
sele=c("39/0","39/1","35/0","35/1","31/0","31/1")
jpeg(file="F:\\projecte\\nens
representatius(model.lme6).jpg",width=768,height=768)
plot(aa2[aa2[,2]%in%sele,],layout=c(2,3,1))
dev.off()

# Fem els gràfics de tots nens per posar-los a l'annex.
sele=c("1/0","1/1","4/0","4/1","8/0","8/1","9/0","9/1","14/0","14/1","15/0","15/1","1
6/0","16/1","24/0","24/1","26/0","26/1","27/0","27/1","31/0","31/1","34/0","34/1","3
5/0","35/1","36/0","36/1","39/0","39/1","42/0","42/1","43/0","43/1","46/0","46/1","4
7/0","47/1","52/0","52/1","54/0","54/1","62/0","62/1","64/0","64/1","65/0","65/1")
jpeg(file="F:\\projecte\\nens tots(model.lme6) 1.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()

sele=c("68/0","68/1","77/0","77/1","81/0","81/1","82/0","82/1","87/0","87/1","88/0",
"88/1","90/0","90/1","91/0","91/1","94/0","94/1","95/0","95/1","97/0","97/1","98/0",
"98/1","101/0","101/1","105/0","105/1","107/0","107/1","113/0","113/1","118/0","11
8/1","122/0","122/1","124/0","124/1","126/0","126/1","130/0","130/1","132/0","132/
1","141/0","141/1","145/0","145/1")
jpeg(file="F:\\projecte\\nens tots(model.lme6) 2.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()

sele=c("147/0","147/1","149/0","149/1","151/0","151/1","154/0","154/1","158/0","15
8/1","160/0","160/1","162/0","162/1","168/0","168/1","170/0","170/1","173/0","173/
1","177/0","177/1","178/0","178/1","180/0","180/1","182/0","182/1","185/0","185/1",
"188/0","188/1","191/0","191/1","196/0","196/1","197/0","197/1","202/0","202/1","
204/0","204/1","205/0","205/1","208/0","208/1","217/0","217/1")
jpeg(file="F:\\projecte\\nens tots(model.lme6) 3.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()

```

```
sele=c("220/0","220/1","223/0","223/1","224/0","224/1","225/0","225/1","226/0","226/1","237/0","237/1","241/0","241/1","242/0","242/1","245/0","245/1","246/0","246/1","249/0","249/1","256/0","256/1","263/0","263/1","264/0","264/1","267/0","267/1","278/0","278/1","279/0","279/1","280/0","280/1","285/0","285/1","286/0","286/1","289/0","289/1","290/0","290/1","293/0","293/1","296/0","296/1")
jpeg(file="F:\\projecte\\nens tots(model.lme6) 4.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("298/0","298/1","301/0","301/1","302/0","302/1","303/0","303/1","307/0","307/1","313/0","313/1","316/0","316/1","320/0","320/1","321/0","321/1","327/0","327/1","333/0","333/1","344/0","344/1","346/0","346/1","347/0","347/1","355/0","355/1","358/0","358/1","359/0","359/1","360/0","360/1","366/0","366/1","367/0","367/1","368/0","368/1","371/0","371/1","375/0","375/1","385/0","385/1")
jpeg(file="F:\\projecte\\nens tots(model.lme6) 5.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("386/0","386/1","397/0","397/1","398/0","398/1","400/0","400/1","401/0","401/1","404/0","404/1","407/0","407/1","408/0","408/1","416/0","416/1","417/0","417/1","418/0","418/1","421/0","421/1","439/0","439/1","447/0","447/1","449/0","449/1","453/0","453/1","457/0","457/1","458/0","458/1","459/0","459/1","461/0","461/1","470/0","470/1","471/0","471/1","473/0","473/1","478/0","478/1")
jpeg(file="F:\\projecte\\nens tots(model.lme6) 6.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("485/0","485/1","488/0","488/1","489/0","489/1","494/0","494/1","499/0","499/1","501/0","501/1","511/0","511/1","519/0","519/1","523/0","523/1","524/0","524/1","526/0","526/1","529/0","529/1","533/0","533/1","538/0","538/1","539/0","539/1","542/0","542/1","549/0","549/1","557/0","557/1","561/0","561/1","565/0","565/1","571/0","571/1","574/0","574/1","575/0","575/1","581/0","581/1")
jpeg(file="F:\\projecte\\nens tots(model.lme6) 7.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("588/0","588/1","589/0","589/1","591/0","591/1","595/0","595/1","606/0","606/1","608/0","608/1","610/0","610/1","630/0","630/1","641/0","641/1","644/0","644/1","646/0","646/1","647/0","647/1","654/0","654/1","656/0","656/1","665/0","665/1","671/0","671/1","674/0","674/1","675/0","675/1","676/0","676/1","681/0","681/1","682/0","682/1","685/0","685/1","690/0","690/1","697/0","697/1")
jpeg(file="F:\\projecte\\nens tots(model.lme6) 8.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("699/0","699/1","701/0","701/1","702/0","702/1","712/0","712/1","713/0","713/1","714/0","714/1","717/0","717/1","718/0","718/1","719/0","719/1","720/0","720/1","721/0","721/1","728/0","728/1","735/0","735/1","738/0","738/1","739/0","739/1","741/0","741/1","742/0","742/1","746/0","746/1","748/0","748/1","753/0","753/1","759/0","759/1","760/0","760/1","761/0","761/1","762/0","762/1")
```

```
jpeg(file="F:\\projecte\\nens tots(model.lme6) 9.jpg",width=768,height=1280)
plot(aa2[aa2[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
aa=augPred(model.lme35,level=0:2)
sele=c("1/0","1/1","4/0","4/1","8/0","8/1","9/0","9/1","14/0","14/1","15/0","15/1","16/0","16/1","24/0","24/1","26/0","26/1","27/0","27/1","31/0","31/1","34/0","34/1","35/0","35/1","36/0","36/1","39/0","39/1","42/0","42/1","43/0","43/1","46/0","46/1","47/0","47/1","52/0","52/1","54/0","54/1","60/0","60/1","62/0","62/1","64/0","64/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 1.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("65/0","65/1","68/0","68/1","74/0","74/1","77/0","77/1","81/0","81/1","82/0","82/1","87/0","87/1","88/0","88/1","90/0","90/1","91/0","91/1","94/0","94/1","95/0","95/1","97/0","97/1","98/0","98/1","101/0","101/1","105/0","105/1","107/0","107/1","113/0","113/1","114/0","114/1","118/0","118/1","122/0","122/1","124/0","124/1","126/0","126/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 2.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("130/0","130/1","132/0","132/1","141/0","141/1","145/0","145/1","147/0","147/1","149/0","149/1","151/0","151/1","154/0","154/1","158/0","158/1","160/0","160/1","162/0","162/1","168/0","168/1","170/0","170/1","173/0","173/1","177/0","177/1","178/0","178/1","180/0","180/1","182/0","182/1","185/0","185/1","188/0","188/1","191/0","191/1","196/0","196/1","197/0","197/1","202/0","202/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 3.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("204/0","204/1","205/0","205/1","208/0","208/1","217/0","217/1","220/0","220/1","223/0","223/1","224/0","224/1","225/0","225/1","226/0","226/1","237/0","237/1","241/0","241/1","242/0","242/1","245/0","245/1","246/0","246/1","249/0","249/1","256/0","256/1","263/0","263/1","264/0","264/1","267/0","267/1","278/0","278/1","279/0","279/1","280/0","280/1","285/0","285/1","286/0","286/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 4.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("289/0","289/1","290/0","290/1","293/0","293/1","296/0","296/1","298/0","298/1","301/0","301/1","302/0","302/1","303/0","303/1","307/0","307/1","313/0","313/1","316/0","316/1","320/0","320/1","321/0","321/1","327/0","327/1","333/0","333/1","344/0","344/1","346/0","346/1","347/0","347/1","350/0","350/1","355/0","355/1","358/0","358/1","359/0","359/1","360/0","360/1","366/0","366/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 5.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("367/0","367/1","368/0","368/1","371/0","371/1","372/0","372/1","375/0","375/1","385/0","385/1","386/0","386/1","388/0","388/1","397/0","397/1","398/0","398/1","400/0","400/1","401/0","401/1","404/0","404/1","407/0","407/1","408/0","408/1","416/0","416/1","417/0","417/1","418/0","418/1","421/0","421/1","439/0","439/1","447/0","447/1","449/0","449/1","453/0","453/1","457/0","457/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 6.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

```
sele=c("458/0","458/1","459/0","459/1","461/0","461/1","470/0","470/1","471/0","471/1","473/0","473/1","478/0","478/1","485/0","485/1","488/0","488/1","489/0","489/1","494/0","494/1","499/0","499/1","501/0","501/1","511/0","511/1","519/0","519/1","523/0","523/1","524/0","524/1","526/0","526/1","529/0","529/1","533/0","533/1","538/0","538/1","539/0","539/1","542/0","542/1","549/0","549/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 7.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

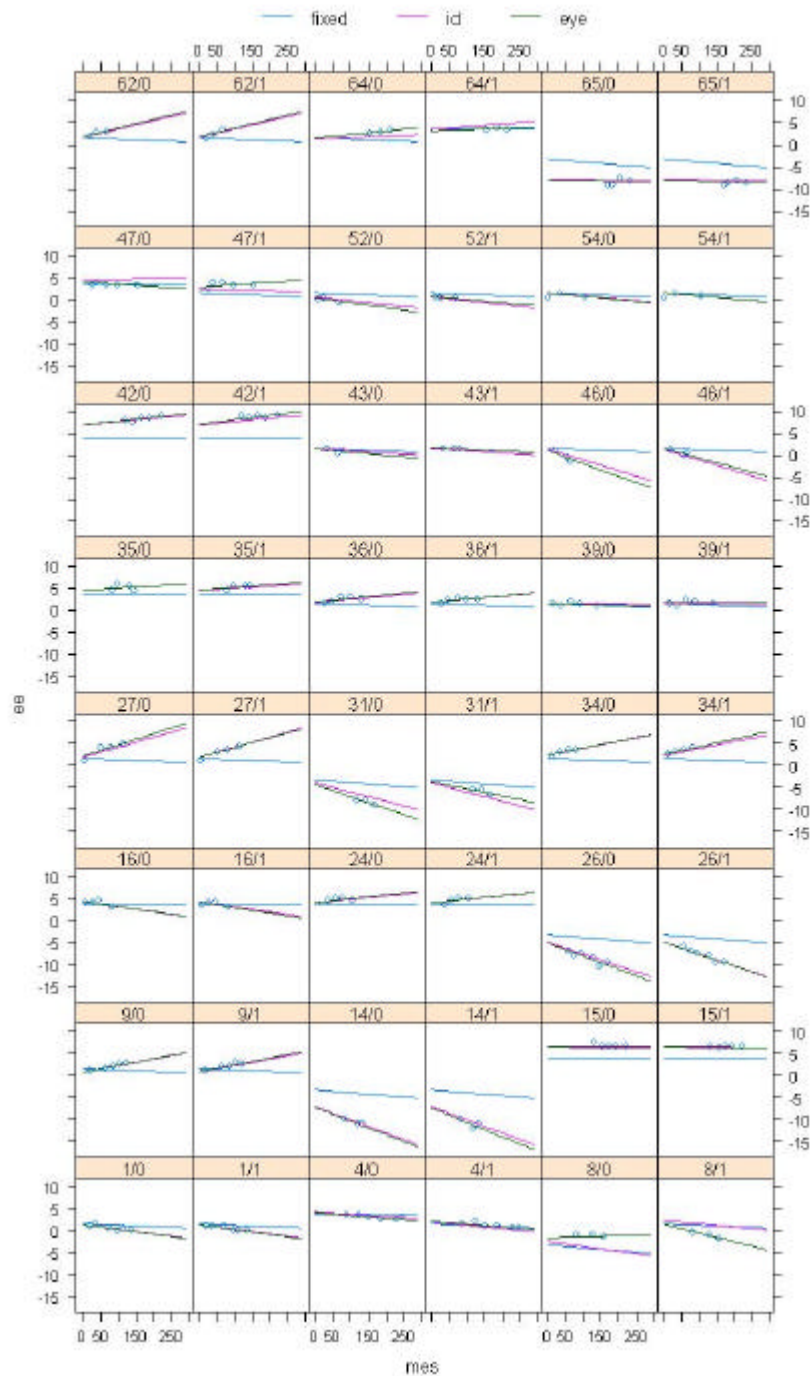
```
sele=c("557/0","557/1","561/0","561/1","565/0","565/1","571/0","571/1","574/0","574/1","575/0","575/1","578/0","578/1","581/0","581/1","588/0","588/1","589/0","589/1","591/0","591/1","595/0","595/1","599/0","599/1","606/0","606/1","608/0","608/1","610/0","610/1","630/0","630/1","633/0","633/1","641/0","641/1","644/0","644/1","646/0","646/1","647/0","647/1","654/0","654/1","656/0","656/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 8.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

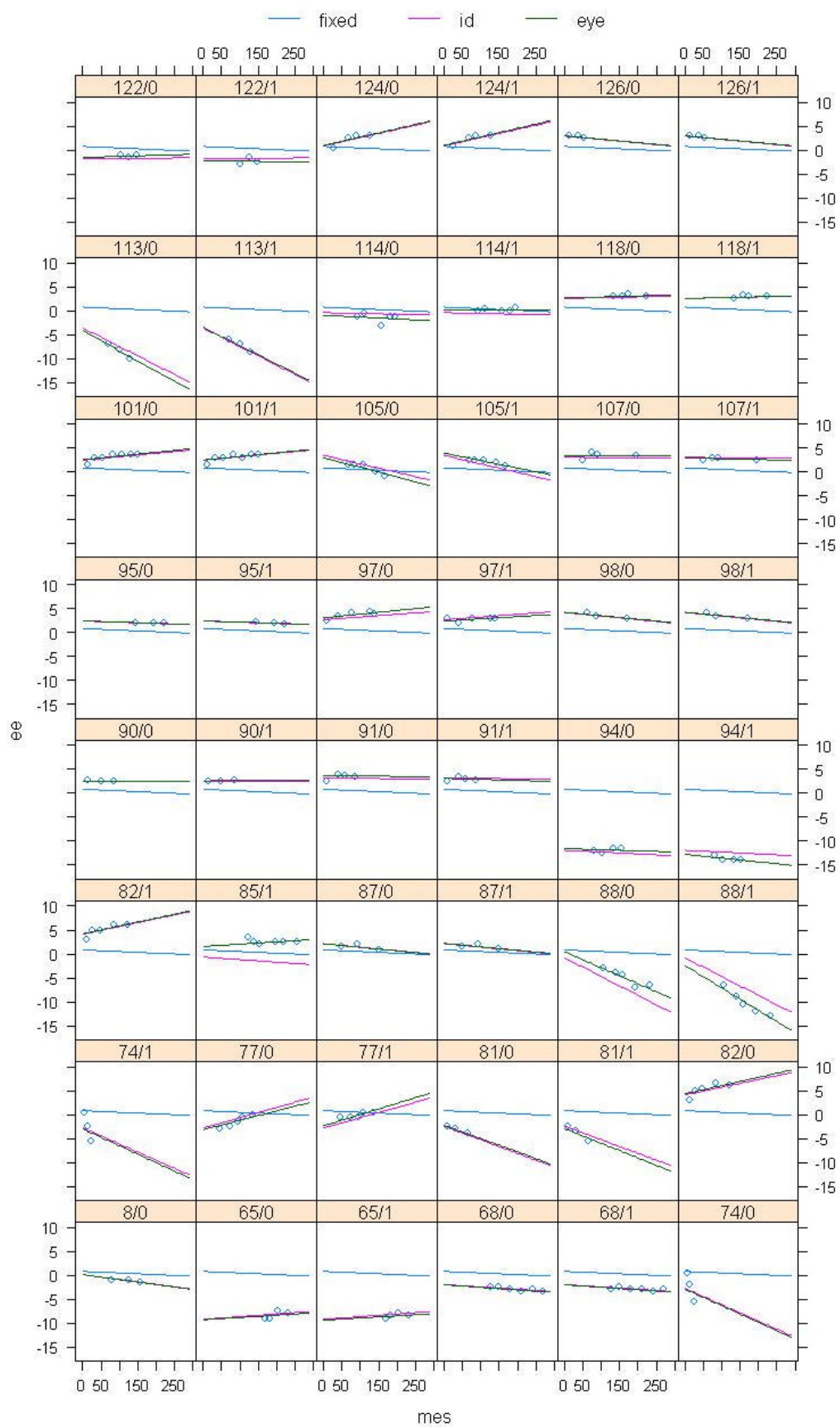
```
sele=c("665/0","665/1","671/0","671/1","672/0","672/1","674/0","674/1","675/0","675/1","676/0","676/1","681/0","681/1","682/0","682/1","685/0","685/1","690/0","690/1","697/0","697/1","699/0","699/1","701/0","701/1","702/0","702/1","703/0","703/1","712/0","712/1","713/0","713/1","714/0","714/1","717/0","717/1","718/0","718/1","719/0","719/1","720/0","720/1","721/0","721/1","728/0","728/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 9.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(6,8,1))
dev.off()
```

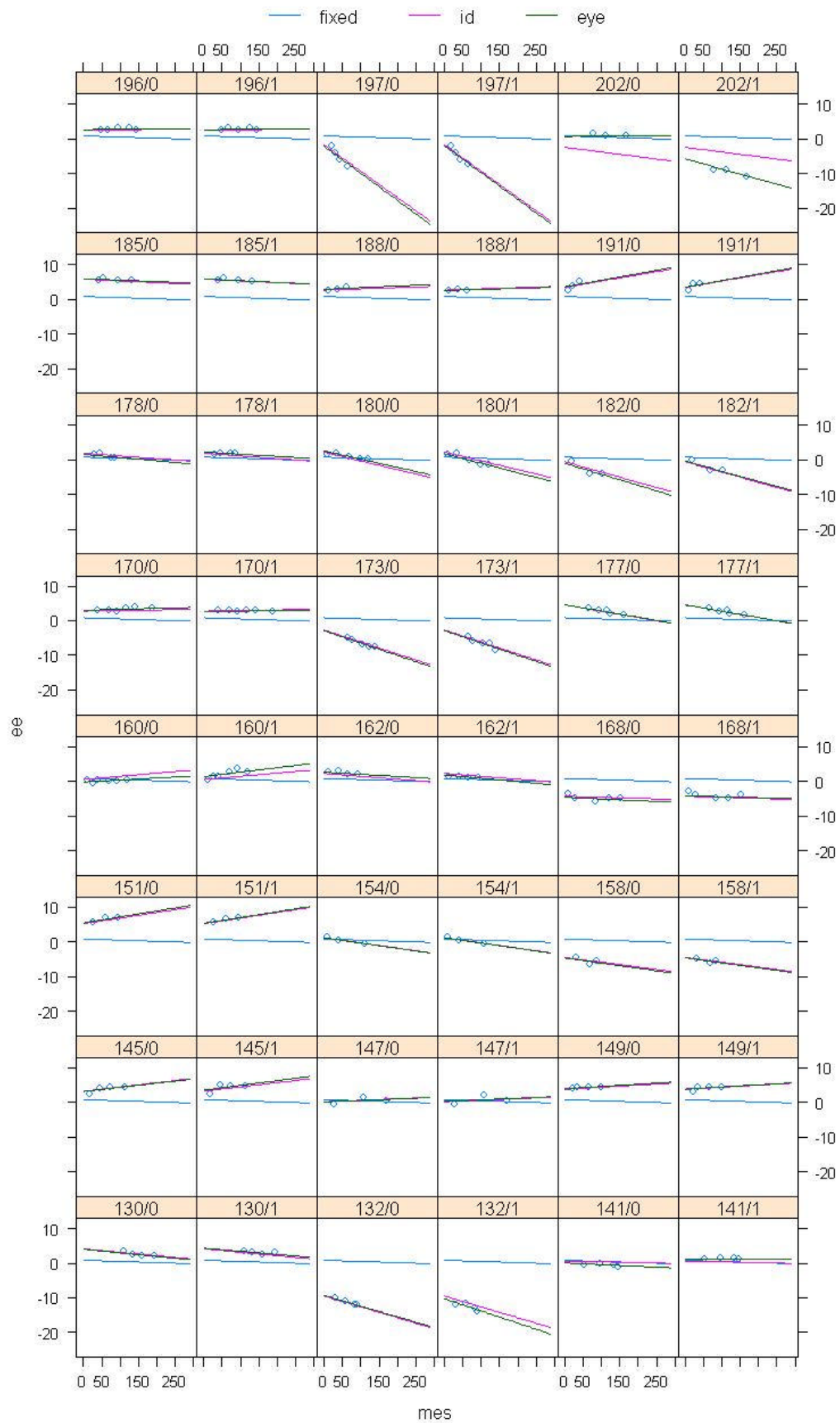
```
sele=c("730/0","730/1","735/0","735/1","738/0","738/1","739/0","739/1","741/0","741/1","742/0","742/1","746/0","746/1","748/0","748/1","753/0","753/1","759/0","759/1","760/0","760/1","761/0","761/1","762/0","762/1")
jpeg(file="F:\\projecte\\nens tots(model.lme35) 10.jpg",width=768,height=1280)
plot(aa[aa[,2]%in%sele,],layout=c(4,6,1))
dev.off()
```

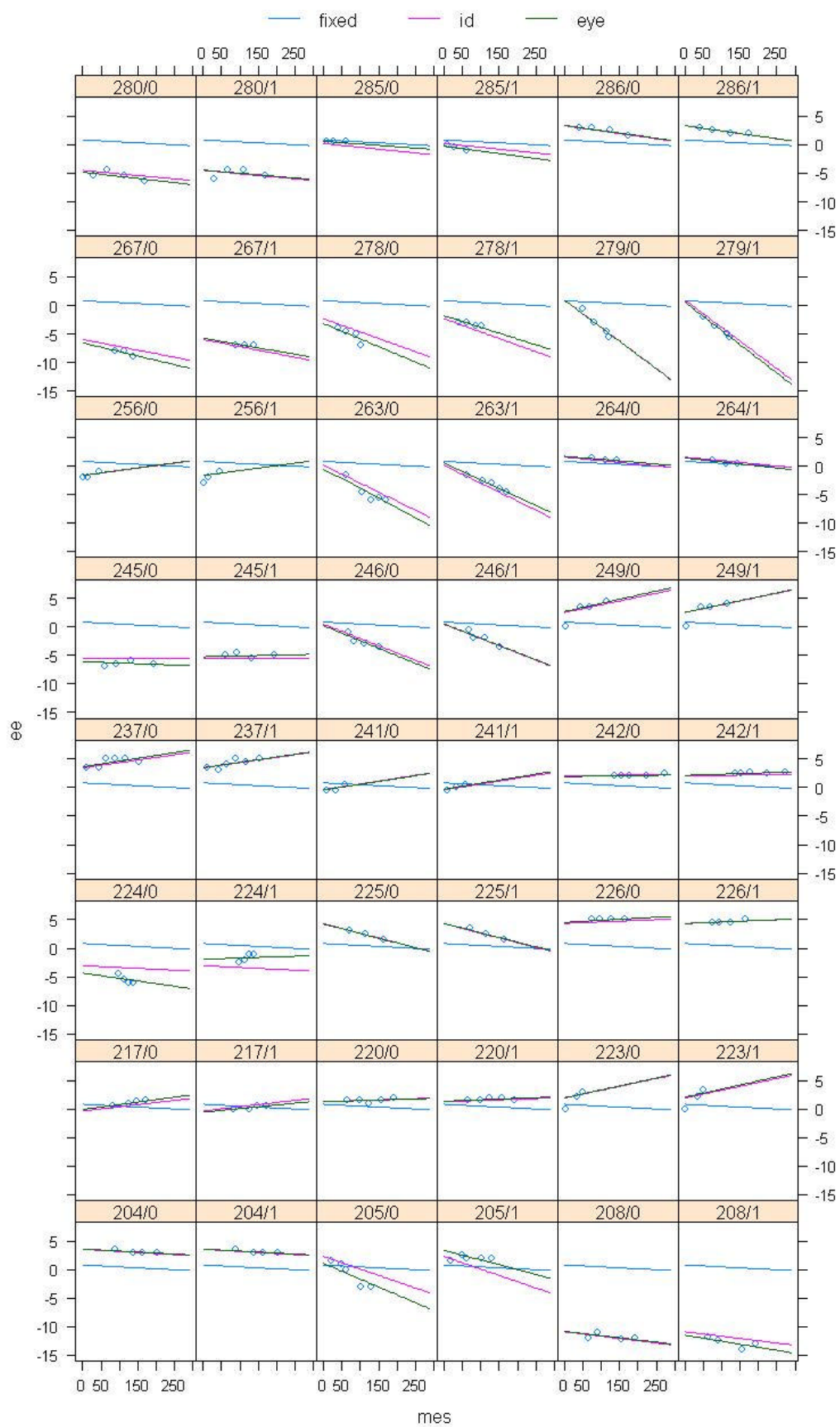
7.3. Modelització dels individus de la mostra

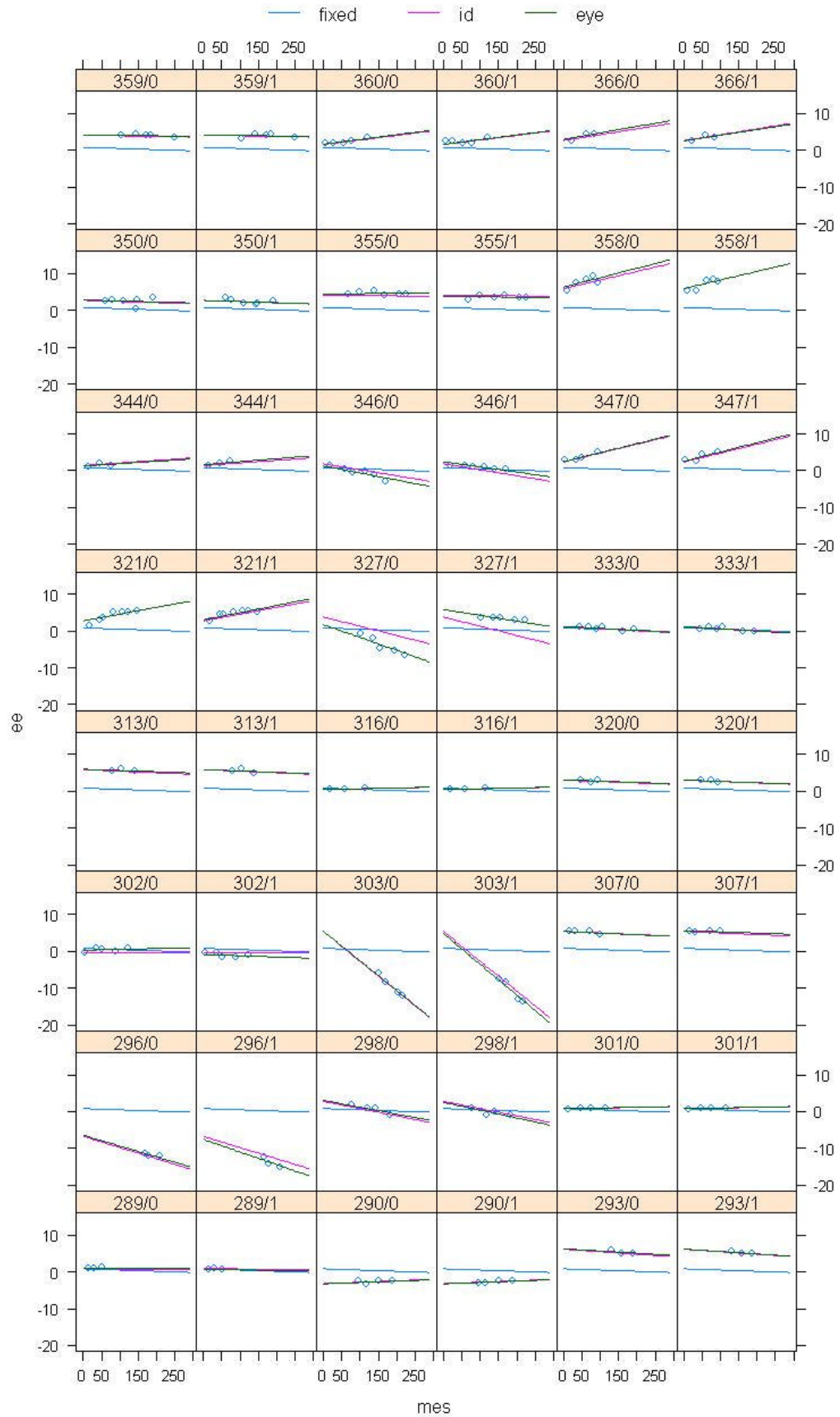
MODEL	EQUACIÓ DEL MODEL	NIVELL NEN	NIVELL ULL	RESIDUS
Model lineal mixt 5	$ee_{ijk} = (0.8356 + b_{0i} + b_{0j}) + (-0.0033 + b_{1i} + b_{1j})mes_{ijk} + e_{ijk}$	$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 3.7274 & 0 \\ 0 & 0.0176 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{0j} \\ b_{1j} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.7515 & 0 \\ 0 & 0.0035 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 0.996 \\ 0.996 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (4.2328)^2 (0.1043 + mes^{-0.8294}))$

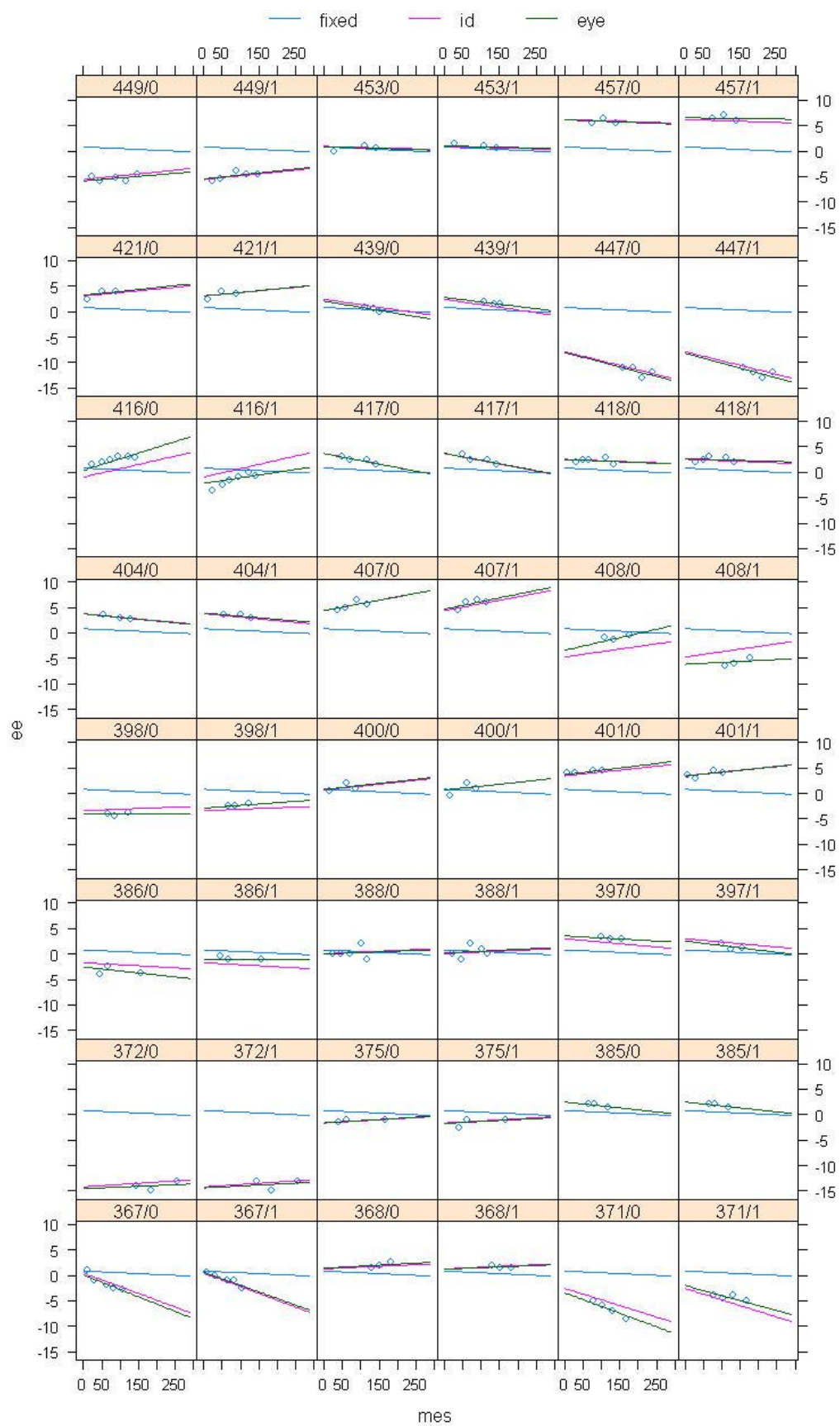


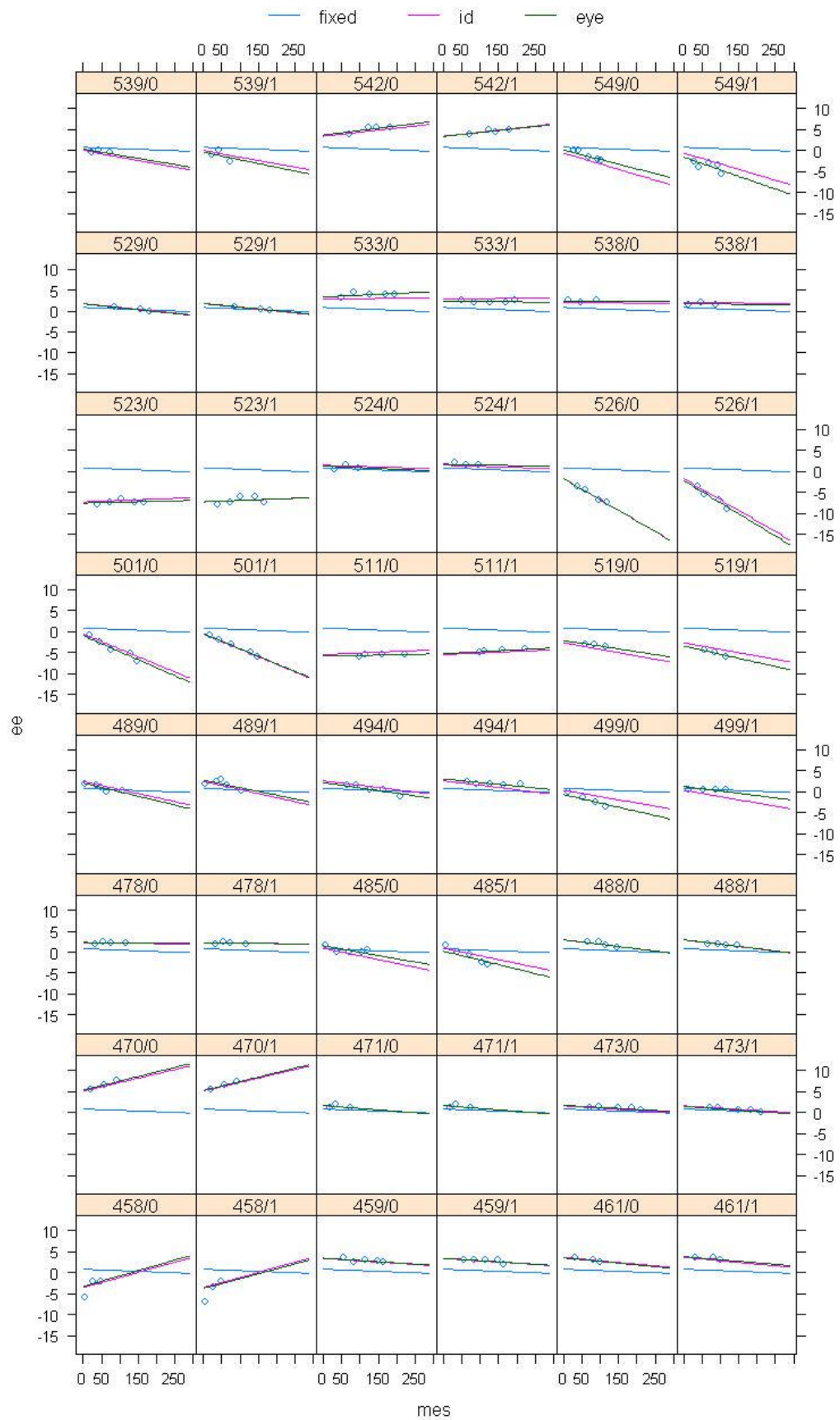


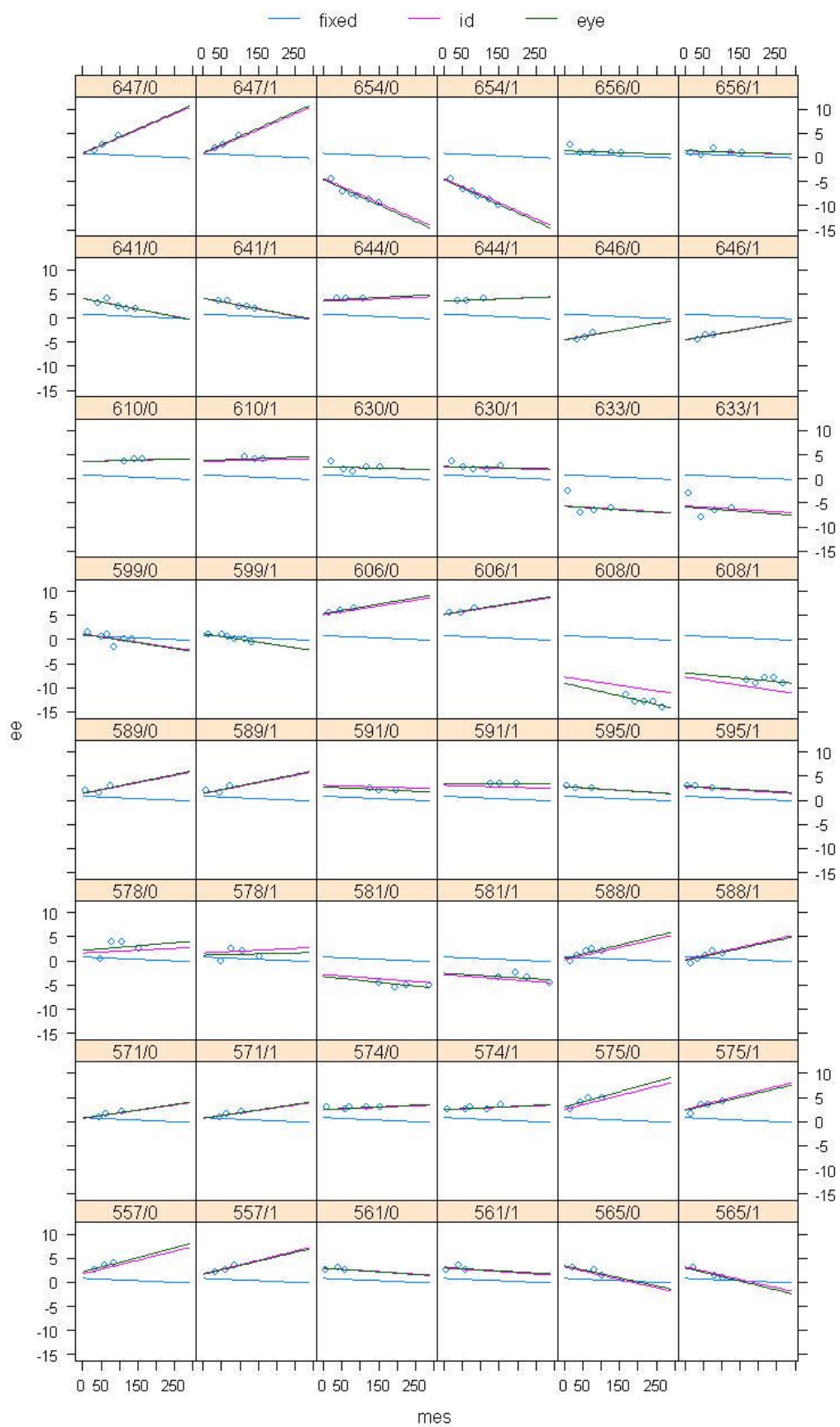


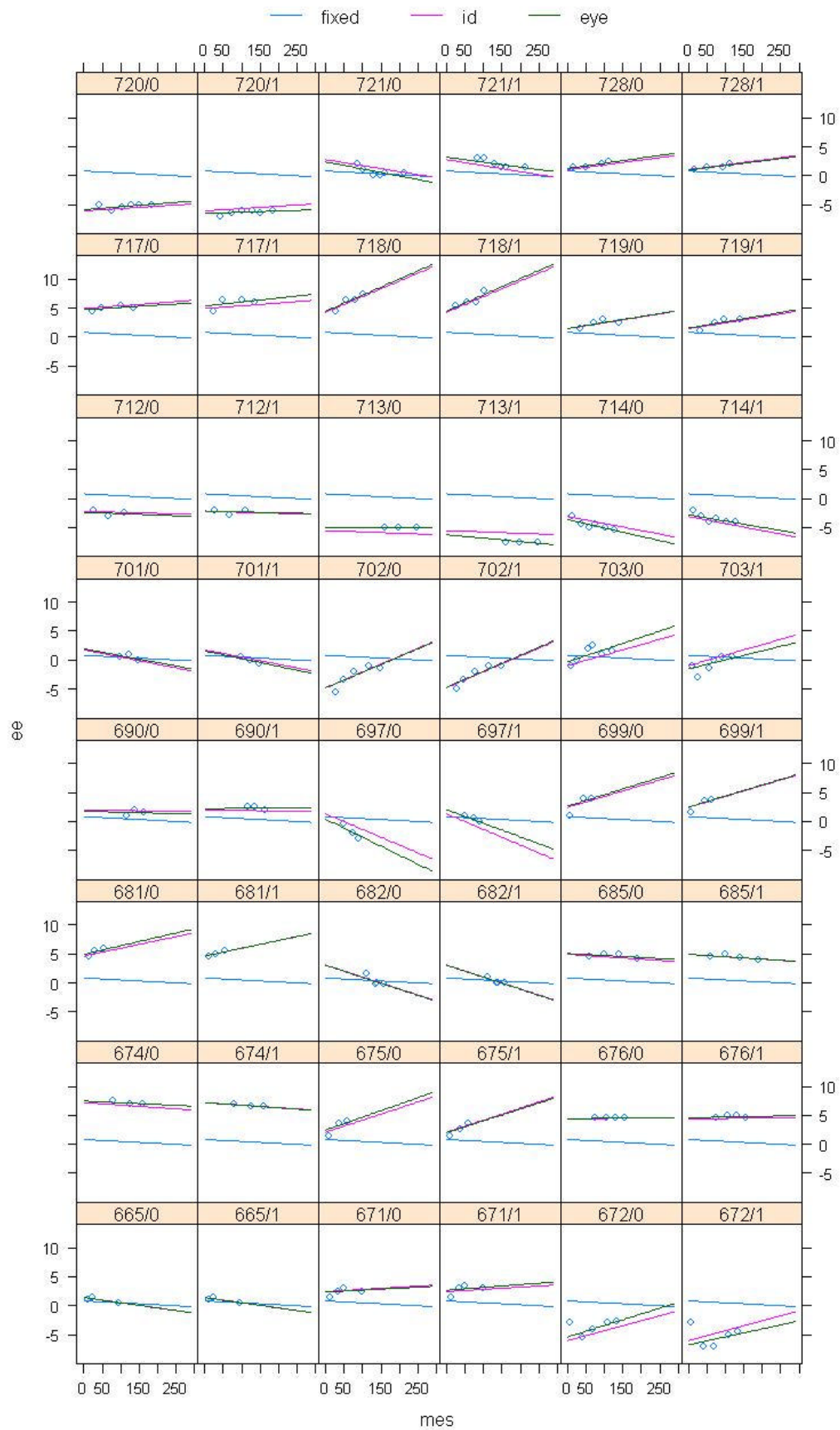


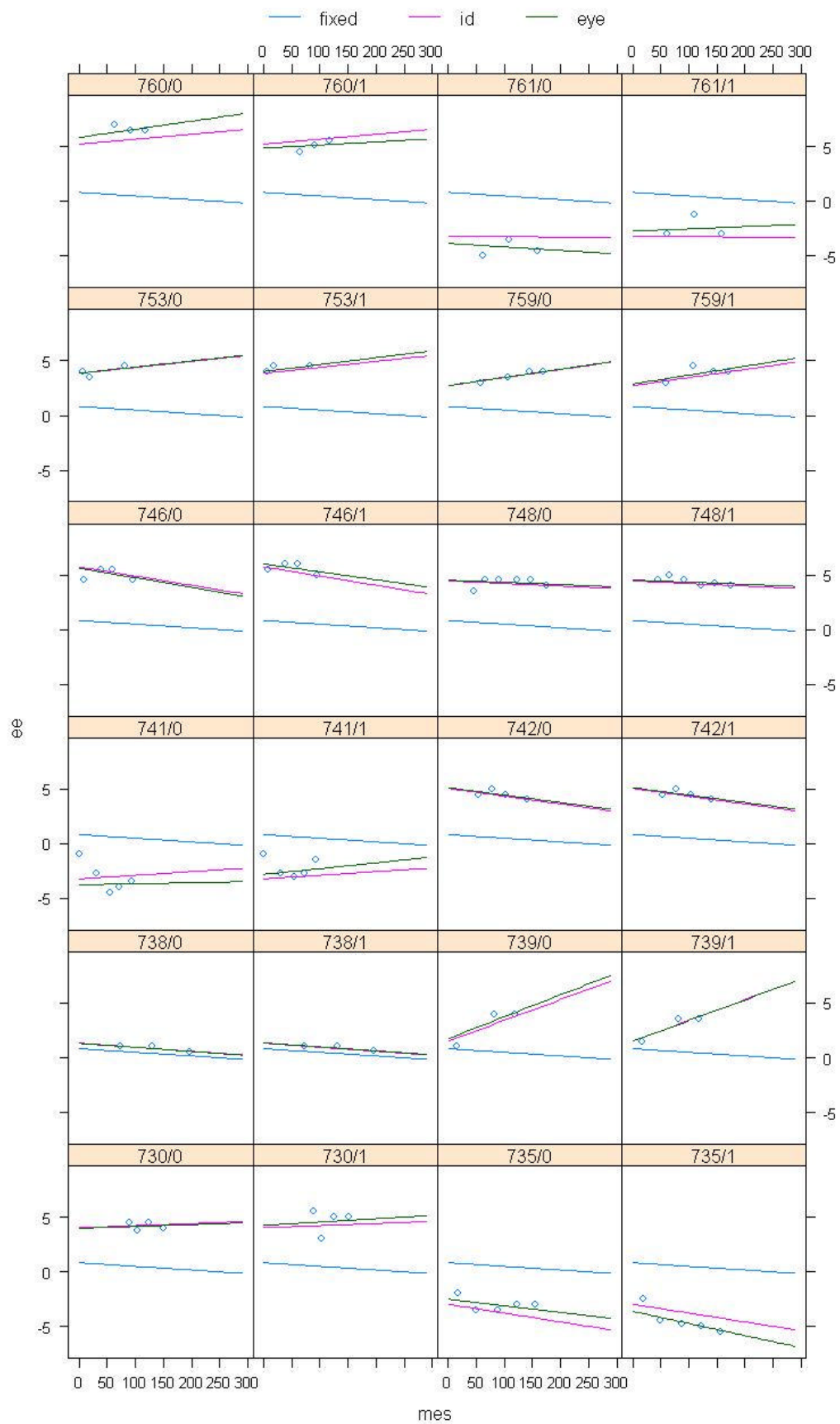












7.4. Modelització dels individus de la mostra sense atípics

MODEL	EQUACIÓ DEL MODEL	NIVELL NEN	NIVELL ULL	RESIDUS
Model lineal mixt sense atípics 3	$e_{ijk} = (-3.227 + 4.787 I_c + 6.974 I_s + b_{ij} + b_{ij}) +$ $+ (-0.006 + 0.003 I_c + 0.006 I_s + b_{ij} + b_{ij}) me_{ijk} +$ $+ e_{ijk}$	$\begin{pmatrix} b_{ij} \\ b_{ij} \end{pmatrix} \sim N(0, V^{(1)} C^{(1)} V^{(1)})$ $V^{(1)} = \begin{pmatrix} 1.6821 & 0 \\ 0 & 0.0178 \end{pmatrix}$ $C^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} b_{ij} \\ b_{ij} \end{pmatrix} \sim N(0, V^{(2)} C^{(2)} V^{(2)})$ $V^{(2)} = \begin{pmatrix} 0.3679 & 0 \\ 0 & 0.0055 \end{pmatrix}$ $C^{(2)} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$e_{ijk} \sim N(0, (1.8093)^2 (0.2250 + me_{ijk}^{0.6833}))$

